

**FIGURE 1**

GGGGCTTCGGCGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGGATTTACAAAAGGTGCAGGTATG  
AGCAGGTCTGAAGACTAACATTTTGTGAAGTTGTAAACAGAAAACCTGTTAGAAATGTGGTGGT  
TTCAGCAAGGCCTCAGTTTCCTTCCTTCAGCCCTTGTAATTTGGACATCTGCTGCTTTCATATTT  
TCATACATTACTGCAGTAACACTCCACCATATAGACCCGGCTTTACCTTATATCAGTGACACTGG  
TACAGTAGCTCCAGAAAAATGCTTATTTGGGGCAATGCTAAATATTGCGGCAGTTTTATGCATTG  
CTACCATTTATGTTTCGTTATAAGCAAGTTCATGCTCTGAGTCCTGAAGAGAACGTTATCATCAA  
TTAAACAAGGCTGGCCTTGTACTTGGAATACTGAGTTGTTTAGGACTTCTATTTGTGGCAAACCT  
CCAGAAAACAACCCTTTTTGTCTGCACATGTAAGTGGAGCTGTGCTTACCTTTGGTATGGGCTCAT  
TATATATGTTTTGTTTCAGACCATCCTTTCCTACCAAATGCAGCCCAAATCCATGGCAAACAAGTC  
TTCTGGATCAGACTGTTGTTGGTTATCTGGTGTGGAGTAAGTGCACTTAGCATGCTGACTTGCTC  
ATCAGTTTTGCACAGTGGCAATTTTGGGACTGATTTAGAACAGAACTCCATTGGAACCCCGAGG  
ACAAAGGTTATGTGCTTCACATGATCACTACTGCAGCAGAATGGTCTATGTCATTTTCCTTCTTT  
GGTTTTTTCCTGACTTACATTCGTGATTTTCAGAAAATTTCTTTACGGGTGGAAGCCAATTTACA  
TGGATTAACCCTCTATGACACTGCACCTTGCCCTATTAACAATGAACGAACACGGCTACTTTCCA  
GAGATATTTGATGAAAGGATAAAATATTTCTGTAATGATTATGATTCTCAGGGATTGGGGAAAGG  
TTCACAGAAGTTGCTTATTCTTCTCTGAAATTTCAACCACTTAATCAAGGCTGACAGTAACACT  
GATGAATGCTGATAATCAGGAAACATGAAAGAAGCCATTTGATAGATTATTCTAAAGGATATCAT  
CAAGAAGACTATTAAAAACACCTATGCCTATACTTTTTTATCTCAGAAAATAAAGTCAAAGACT  
ATG

## **FIGURE 2**

<subunit 1 of 1, 266 aa, 1 stop

<MW: 29766, pI: 8.39, NX(S/T): 0

MWWFQQGLSFLPSALVIWTSAAFIIFS YITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNIAAV  
LCIATIYVRYKQVHALSPEENVIIKLNKAGLV LGILSCLGLSIVANFQKTTLFAAHVSGAVLTFG  
MGSLYMFVQTILSYQMOPKIHGKQVFWIRLLLVIWCGVSALSMLTCSSVLHSGNFGTDLEQKLHW  
NPEDKGYVLHMITTAAEWSMSFSFFGFFLT YIRDFQKISLRVEANLHGLTLYDTAPCPINNERTR  
LLSRDI

### **Important features:**

#### **Type II transmembrane domain:**

amino acids 13-33

#### **Other Transmembrane domains:**

amino acids 54-73, 94-113, 160-180, 122-141

#### **N-myristoylation sites.**

amino acids 57-63, 95-101, 99-105, 124-130, 183-189

**FIGURE 3**

CGGACGCGTGGGCGGACGCGTGGGGGAGAGCCGCGAGTCCCGGCTGCAGCACCTGGGAGAAGGCAG  
ACCGTGTGAGGGGGCCTGTGGCCCCAGCGTGCTGTGGCCTCGGGGAGTGGGAAGTGGAGGCAGGA  
GCCTTCCTTACACTTCGCC**ATG**AGTTTCCTCATCGACTCCAGCATCATGATTACCTCCCAGATAC  
TATTTTTTGGATTGGGTGGCTTTTCTTCATGCGCCAATTGTTTAAAGACTATGAGATACGTCAG  
TATGTTGTACAGGTGATCTTCTCCGTGACGTTTGCATTTTCTTGACCATGTTTGAGCTCATCAT  
CTTTGAAATCTTAGGAGTATTGAATAGCAGCTCCCGTTATTTTCACTGGAAATGAACCTGTGTG  
TAATTCTGCTGATCCTGGTTTTTCATGGTGCCTTTTTACATTGGCTATTTTATTGTGAGCAATATC  
CGACTACTGCATAAACAACGACTGCTTTTTTCCGTGTCTCTTATGGCTGACCTTTATGTATTTCTT  
CTGGAACTAGGAGATCCCTTTCCCATTCCTCAGCCCAAAACATGGGATCTTATCCATAGAACAGC  
TCATCAGCCGGGTGGTGTGATTGGAGTGAATCTCATGGCTCTTCTTTCTGGATTTGGTGTGTC  
AACTGCCCATACACTTACATGTCTTACTTCCTCAGGAATGTGACTGACACGGATATTCTAGCCCT  
GGAACGGCGACTGCTGCAAACCATGGATATGATCATAAGCAAAAAGAAAAGGATGGCAATGGCAC  
GGAGAACAATGTTCCAGAAGGGGGAAGTGCATAACAAACCATCAGGTTTCTGGGGAATGATAAAA  
AGTGTTACCACTTCAGCATCAGGAAGTGAAAATCTTACTCTTATTCAACAGGAAGTGGATGCTTT  
GGAAGAATTAAGCAGGCAGCTTTTTCTGGAAACAGCTGATCTATATGCTACCAAGGAGAGAATAG  
AATACTCCAAAACCTTCAAGGGGAAATATTTTAATTTTCTTGGTTACTTTTTCTCTATTTACTGT  
GTTTGGAAAATTTTCATGGCTACCATCAATATTGTTTTTGATCGAGTTGGGAAAACGGATCCTGT  
CACAAGAGGCATTGAGATCACTGTGAATTATCTGGGAATCCAATTTGATGTGAAGTTTTGGTCCC  
AACACATTTCTTCATTCTTGTGGAATAATCATCGTCACATCCATCAGAGGATTGCTGATCACT  
CTTACCAAGTTCTTTTATGCCATCTCTAGCAGTAAGTCCTCCAATGTCATTGTCCTGCTATTAGC  
ACAGATAATGGGCATGTACTTTGTCTCCTCTGTGCTGCTGATCCGAATGAGTATGCCTTTAGAAT  
ACCGCACCATTAATCACTGAAGTCCTTGGAGAACTGCAGTTCAACTTCTATCACCGTTGGTTTGAT  
GTGATCTTCCTGGTCAGCGCTCTCTCTAGCATACTCTTCCTCTATTTGGCTCACAAACAGGCACC  
AGAGAAGCAAATGGCACCT**TGA**ACTTAAGCCTACTACAGACTGTTAGAGGCCAGTGGTTTCAAAA  
TTTAGATATAAGAGGGGGGAAAAATGGAACCAGGGCCTGACATTTTATAAACAAACAAAATGCTA  
TGGTAGCATTTTTTACCTTCATAGCATACTCCTTCCCCGTCAGGTGATACTATGACCATGAGTAG  
CATCAGCCAGAACATGAGAGGGGAGAACTAACTCAAGACAATACTCAGCAGAGAGCATCCCGTGTG  
GATATGAGGCTGGTGTAGAGGCGGAGAGGAGCCAAGAACTAAAGGTGAAAAATACACTGGAACCT  
CTGGGGCAAGACATGTCTATGGTAGCTGAGCCAAACACGTAGGATTTCCGTTTTTAAGGTTACAT  
GAAAAGGTTATAGCTTTGCCTTGAGATTGACTCATTAAAATCAGAGACTGTAAACAAAAA  
AAAAAAGGCGGCGCGACTCTAGAGTCGACCTGCAGAAGCTTGGCCGCCATGGCCCAAC  
TTGTTTATTGCAGCTTATAATG

## **FIGURE 4**

MSFLIDSSIMITSQILFFGFGWLFFMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEILGV  
LNSSSRYPHWMNLCVILLILVFMVPFYIGYFIVSNIRLLHKQRLLFSCLLWLTFMYFFWKLGBP  
FPILSPKHGILSIEQLISRVGIVGVTLMALLSGFGAVNCPYTYMSYFLRNVTDTDILALERLLQ  
TMDMIISKKKRMAMARRTMFQKGEVHNKPSGFWGMIKSVTTSASGSENLTLIQQEVDAL EELSRQ  
LFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVFDRVGKTDPVTRGIEI  
TVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKEFFYAISSSKSSNVIVLLLAQIMGY  
FVSSVLLIRMSMPLEYRTIITEVLGELQFNFYHRWFDVIFLVSALSSILFLYLAHKQAPEKQMAP

**Important features:**

**Signal peptide:**

amino acids 1-23

**Potential transmembrane domains:**

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,  
425-444

**N-glycosylation sites.**

amino acids 67-70, 180-183 and 243-246

**Eukaryotic cobalamin-binding proteins**

amino acids 151-160



**FIGURE 5**

AGCAGGGAAATCCGGATGTCTCGGTTATGAAGTGGAGCAGTGAGTGTGAGCCTCAACATAGTTCC  
 AGAACTCTCCATCCGGACTAGTTATTGAGCATCTGCCTCTCATATCACCAGTGGCCATCTGAGGT  
 GTTTCCTTGGCTCTGAAGGGGTAGGCACG**ATG**GCCAGGTGCTTCAGCCTGGTGTGCTTCTCACT  
 TCCATCTGGACCACGAGGCTCCTGGTCCAAGGCTCTTTGCGTGCAGAAGAGCTTTCCATCCAGGT  
 GTCATGCAGAATTATGGGGATCACCCCTTGTGAGCAAAAAGGCGAACCAGCAGCTGAATTTACAG  
 AAGCTAAGGAGGCCTGTAGGCTGCTGGGACTAAGTTTGGCCGGCAAGGACCAAGTTGAAACAGCC  
 TTGAAAGCTAGCTTTGAAACTTGCAGCTATGGCTGGGTGGAGATGGATTTCGTGGTCATCTCTAG  
 GATTAGCCCAAACCCCAAGTGTGGGAAAAATGGGGTGGGTGTCCTGATTTGGAAGGTTCCAGTGA  
 GCCGACAGTTTGCAGCCTATTGTTACAACCTCATCTGATACTTGGACTAACTCGTGCATTCCAGAA  
 ATTATCACCACCAAAGATCCCATATTCAACACTCAAACCTGCAACACAAACAACAGAATTTATTGT  
 CAGTGACAGTACCTACTCGGTGGCATCCCCTTACTCTACAATACCTGCCCCTACTACTACTCCTC  
 CTGCTCCAGCTTCCACTTCTATTCCACGGAGAAAAAAATTGATTTGTGTACAGAAAGTTTTTATG  
 GAACTAGCACCATGTCTACAGAACTGAACCATTTGTTGAAAATAAAGCAGCATTCAAGAATGA  
 AGCTGCTGGGTTTGGAGGTGTCCCCACGGCTCTGCTAGTGCTTGCTCTCCTCTTCTTTGGTGCTG  
 CAGCTGGTCTTGGATTTTGCTATGTCAAAAGGTATGTGAAGGCCTTCCCTTTTACAAACAAGAAT  
 CAGCAGAAGGAAATGATCGAAACCAAAGTAGTAAAGGAGGAGAAGGCCAATGATAGCAACCCTAA  
 TGAGGAATCAAAGAAAACCTGATAAAAACCCAGAAGAGTCCAAGAGTCCAAGCAAAACTACCGTGC  
 GATGCCTGGAAGCTGAAGTT**TAG**ATGAGACAGAAATGAGGAGACACACCTGAGGCTGGTTTCTTT  
 CATGCTCCTTACCCTGCCCCAGCTGGGGAAATCAAAGGGCCAAAGAACCAAAGAAGAAAGTCCA  
 CCCTTGGTTCCCTAACTGGAATCAGCTCAGGACTGCCATTGGACTATGGAGTGCACCAAAGAGAAT  
 GCCCTTCTCCTTATTGTAACCCTGTCTGGATCCTATCCTCCTACCTCCAAAGCTTCCCACGGCCT  
 TTCTAGCCTGGCTATGTCCTAATAATATCCCCTGGGAGAAAGGAGTTTTCGAAAGTGCAAGGAC  
 CTAACATCTCATCAGTATCCAGTGGTAAAAAGGCCTCCTGGCTGTCTGAGGCTAGGTGGGTG  
 AAAGCCAAGGAGTCACTGAGACCAAGGCTTTCTCTACTGATTCCGCAGCTCAGACCCTTTCTTCA  
 GCTCTGAAAGAGAAACACGTATCCACCTGACATGTCTTCTGAGCCCGGTAAGAGCAAAAGAAT  
 GGCAGAAAAGTTTAGCCCCTGAAAGCCATGGAGATTCTCATAACTTGAGACCTAATCTCTGTAAA  
 GCTAAAATAAAGAAATAGAACAAGGCTGAGGATACGACAGTACACTGTCAGCAGGGACTGTAAAC  
 ACAGACAGGGTCAAAGTGTTTTCTCTGAACACATTGAGTTGGAATCACTGTTTAGAACACACACA  
 CTTACTTTTTCTGGTCTCTACCACTGCTGATATTTTCTCTAGGAAATATACTTTTACAAGTAACA  
 AAAATAAAAACCTCTTATAAATTTCTATTTTTATCTGAGTTACAGAAATGATTACTAAGGAAGATT  
 ACTCAGTAATTTGTTTAAAAAGTAATAAAATTCAACAAACATTTGCTGAATAGCTACTATATGTC  
 AAGTGCTGTGCAAGGTATTACACTCTGTAATTGAATATTATTCCTCAAAAAATTGCACATAGTAG  
 AACGCTATCTGGGAAGCTATTTTTTTTCAGTTTTGATATTTCTAGCTTATCTACTTCCAACTAAT  
 TTTTATTTTTGCTGAGACTAATCTTATTCATTTTCTCTAATATGGCAACCATTATAACCTTAATT  
 TATTATTAACATACCTAAGAAGTACATTGTTACCTCTATATACCAAAGCACATTTTAAAAGTGCC  
 ATTAACAAATGTATCACTAGCCCTCCTTTTTTCCAACAAGAAGGGACTGAGAGATGCAGAAATATT  
 TGTGACAAAAAATTAAAGCATTTAGAAAACCTT

## **FIGURE 6**

MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLG  
LSLAGKDQVETALKASFETCSYGWVGDFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYN  
SSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTTTPPAPASTSIPR  
RKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTALLVLALLFFGAAAGLGFCYVK  
RYVKAFFFTNKNQQKEMIETKVVKEEKANDSNPNEESKKTDKNPEESKSPSKTTVRCLEAEV

**Signal sequence:**

amino acids 1-16

**Transmembrane domain:**

amino acids 235-254

**N-glycosylation site.**

amino acids 53-57, 130-134, 289-293

**Casein kinase II phosphorylation site.**

amino acids 145-149, 214-218

**Tyrosine kinase phosphorylation site.**

amino acids 79-88

**N-myristoylation site.**

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

**FIGURE 7**

CGCCGCGCTCCCGCACCCGCGGCCCCGCCACCGCGCCGCTCCCGCATCTGCACCCGCAGCCCGGC  
GGCCTCCCGGCGGGAGCGAGCAGATCCAGTCCGGCCCCGCAGCGAACTCGGTCCAGTCGGGGCGG  
CGGCTGCGGGGCGCAGAGCGGAGATGCAGCGGCTTGGGGCCACCCTGCTGTGCTGCTGCTGGCGG  
CGGCGGTCCCCACGGCCCCCGCGCCCGCTCCGACGGCGACCTCGGCTCCAGTCAAGCCCGGCCCG  
GCTCTCAGCTACCCGCAGGAGGAGGCCACCCTCAATGAGATGTTCCGCGAGGTTGAGGAACTGAT  
GGAGGACACGCAGCACAAATTGCGCAGCGCGGTGGAAGAGATGGAGGCAGAAGAAGCTGCTGCTA  
AAGCATCATCAGAAGTGAACCTGGCAAACCTTACCTCCCAGCTATCACAATGAGACCAACACAGAC  
ACGAAGGTTGGAAATAATACCATCCATGTGCACCGAGAAATTCACAAGATAACCAACAACCAGAC  
TGGACAAATGGTCTTTTCAGAGACAGTTATCACATCTGTGGGAGACGAAGAAGGCAGAAGGAGCC  
ACGAGTGCATCATCGACGAGGACTGTGGGCCCCAGCATGTACTGCCAGTTTGCCAGCTTCCAGTAC  
ACCTGCCAGCCATGCCGGGGCCAGAGGATGCTCTGCACCCGGGACAGTGAGTGCTGTGGAGACCA  
GCTGTGTGTCTGGGGTCACTGCACCAAAATGGCCACCAGGGGCAGCAATGGGACCATCTGTGACA  
ACCAGAGGGACTGCCAGCCGGGGCTGTGCTGTGCCTTCCAGAGAGGCCTGCTGTTCCCTGTGTGC  
ACACCCCTGCCCGTGGAGGGCGAGCTTTGCCATGACCCGCCAGCCGGCTTCTGGACCTCATCAC  
CTGGGAGCTAGAGCCTGATGGAGCCTTGGACCGATGCCCTTGTGCCAGTGGCCTCCTCTGCCAGC  
CCCACAGCCACAGCCTGGTGTATGTGTGCAAGCCGACCTTCGTGGGGAGCCGTGACCAAGATGGG  
GAGATCCTGCTGCCCAGAGAGGTCCCCGATGAGTATGAAGTTGGCAGCTTCATGGAGGAGGTGCG  
CCAGGAGCTGGAGGACCTGGAGAGGAGCCTGACTGAAGAGATGGCGCTGGGGGAGCCTGCGGCTG  
CCGCCGCTGCACTGCTGGGAGGGGAAGAGATTTAGATCTGGACCAGGCTGTGGGTAGATGTGCAA  
TAGAAATAGCTAATTTATTTCCCCAGGTGTGTGCTTTAGGCGTGGGCTGACCAGGCTTCTTCCCTA  
CATCTTCTTCCCAGTAAGTTTCCCCTCTGGCTTGACAGCATGAGGTGTTGTGCATTTGTTTCAGCT  
CCCCAGGCTGTTCTCCAGGCTTCACAGTCTGGTGCTTGGGAGAGTCAGGCAGGGTTAAACTGCA  
GGAGCAGTTTGCCACCCCTGTCCAGATTATTGGCTGCTTTGCCCTCTACCAGTTGGCAGACAGCCG  
TTTGTTCTACATGGCTTTGATAATTGTTTGAGGGGAGGAGATGGAAACAATGTGGAGTCTCCCTC  
TGATTGGTTTTTGGGGAAATGTGGAGAAGAGTGCCCTGCTTTGCAAACATCAACCTGGCAAAAATG  
CAACAAATGAATTTTCCACGCAGTTCTTTCCATGGGCATAGGTAAGCTGTGCCCTTCAGCTGTTGC  
AGATGAAATGTTCTGTTTCACCCTGCATTACATGTGTTTATTCATCCAGCAGTGTTGCTCAGCTCC  
TACCTCTGTGCCAGGGCAGCATTTTCATATCCAAGATCAATTCCTCTCTCAGCACAGCCTGGGG  
AGGGGGTCATTGTTCTCCTCGTCCATCAGGGATCTCAGAGGCTCAGAGACTGCAAGCTGCTTGCC  
CAAGTCACACAGCTAGTGAAGACCAGAGCAGTTTCATCTGGTTGTGACTCTAAGCTCAGTGCTCT  
CTCCACTACCCACACCAGCCTTGGTGCCACCAAAAGTGCTCCCCAAAAGGAAGGAGAATGGGAT  
TTTTCTTGAGGCATGCACATCTGGAATTAAGGTCAAACCTAATTCTCACATCCCTCTAAAAGTAAA  
CTACTGTTAGGAACAGCAGTGTTCTCACAGTGTGGGGCAGCCGTCCTTCTAATGAAGACAATGAT  
ATTGACACTGTCCCTCTTTGGCAGTTGCATTAGTAACCTTTGAAAGGTATATGACTGAGCGTAGCA  
TACAGGTTAACTGCAGAAACAGTACTTAGGTAATTGTAGGGCGAGGATTATAAATGAAATTTGC  
AAAATCACTTAGCAGCAACTGAAGACAATTATCAACCACGTGGAGAAAATCAAACCGAGCAGGGC  
TGTGTGAAACATGGTTGTAATATGCGACTGCGAACACTGAACTCTACGCCACTCCACAAATGATG  
TTTTCAGGTGTCATGGACTGTTGCCACCATGTATTTCATCCAGAGTTCTTAAAGTTTAAAGTTGCA  
CATGATTGTATAAGCATGCTTTCTTTGAGTTTAAATTATGTATAAACATAAGTTGCATTTAGAA  
ATCAAGCATAAATCACTTCACTGCAAAAAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 8**

MQRLGATLLCLLLAAAVPTAPAPAPTATSAPVKPGPALSYPQEEATLNEMFREVEELMEDTQHKL  
RSAVEEMEAEAAAKASSEVNLANLPPSYHNETNTDTKVGNNTIHVHREIHKITNNQTGQMVFSE  
TVITSVGDEEGRRSHECIIDEDCGPSMYCQFASFQYTCQPCRGQRMLCTRDSECCGDQLCVWGHC  
TKMATRGSNGTICDNQRDCQPGLCCAFQRGLLFPVCTPLPVEGELCHDPASRLLDLITWELEPDG  
ALDRPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEVGSFMEEVRQELEDLE  
RSLTEEMALGEPAAAAAALLGGEEI

**Signal sequence:**

amino acids 1-19

**N-glycosylation site.**

amino acids 96-100, 106-110, 121-125, 204-208

**Casein kinase II phosphorylation site.**

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316, 327-331

**N-myristoylation site.**

amino acids 202-208, 217-223

**Amidation site.**

amino acids 140-144

**FIGURE 9**

**CGG**ACGCGTGGGCGGACGCGTGGGGGCTGTGAGAAAGTGCCAATAAATACATCATGCAACCCAC  
GGCCACCTTGTGAACTCCTCGTGCCAGGGCTGATGTGCGTCTTCCAGGGCTACTCATCCAAAG  
GCCTAATCCAACGTTCTGTCTTCAATCTGCAAATCTATGGGGTCCTGGGGCTCTTCTGGACCCTT  
AACTGGGTACTGGCCCTGGGCCAATGCGTCCTCGCTGGAGCCTTTGCCTCCTTCTACTGGGCCTT  
CCACAAGCCCCAGGACATCCCTACCTTCCCCTTAATCTCTGCCTTCATCCGCACACTCCGTTACC  
ACACTGGGTCATTGGCATTGAGCCCTCATCCTGACCCTTGTGCAGATAGCCCGGGTCATCTTG  
GAGTATATTGACCACAAGCTCAGAGGAGTGCAGAACCCTGTAGCCCGCTGCATCATGTGCTGTTT  
CAAGTGCTGCCTCTGGTGTCTGGAAAAATTTATCAAGTTCCTAAACCGCAATGCATACATCATGA  
TCGCCATCTACGGGAAGAATTTCTGTGTCTCAGCCAAAAATGCGTTCATGCTACTCATGCGAAAC  
ATTGTCAGGGTGGTCGTCCTGGACAAAGTCACAGACCTGCTGCTGTTCTTTGGGAAGCTGCTGGT  
GGTCGGAGGCGTGGGGGTCCTGTCCTTCTTTTTTTTCTCCGGTCGCATCCCGGGGCTGGGTAAAG  
ACTTTAAGAGCCCCCACCTCAACTATTACTGGCTGCCCATCATGACCTCCATCCTGGGGGCCTAT  
GTCATCGCCAGCGGCTTCTTCAGCGTTTTTCGGCATGTGTGTGGACACGCTCTTCCTCTGCTTCCT  
GGAAGACCTGGAGCGGAACAACGGCTCCCTGGACCGGCCCTACTACATGTCCAAGAGCCTTCTAA  
AGATTCTGGGCAAGAAGAACGAGGCGCCCCCGGACAACAAGAAGAGGAAGAAG**TGA**CAGCTCCGG  
CCCTGATCCAGGACTGCACCCCACCCCCACCGTCCAGCCATCCAACCTCACTTCGCCTTACAGGT  
CTCCATTTTGTGGTAAAAAAAGGTTTTAGGCCAGGCGCCGTGGCTCACGCCTGTAATCCAACACT  
TTGAGAGGCTGAGGCGGGCGGATCACCTGAGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTG  
AAACCTCCGTCTCTATTAAAAATACAAAAATTAGCCGAGAGTGGTGGCATGCACCTGTCATCCCA  
GCTACTCGGGAGGCTGAGGCAGGAGAATCGCTTGAACCCGGGAGGCAGAGGTTGCAGTGAGCCGA  
GATCGCGCCACTGCACTCCAACCTGGGTGACAGACTCTGTCTCCAAAACAAAACAAACAAA  
AAGATTTTATTAAAGATATTTTGTAACTC

## **FIGURE 10**

RTRGRTRGGCEKVPINTSCNPTAHLVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVLGLEFWTL  
NWVLALGQCVLGAFASFYWAFHKPQDIPTFPLISAFIRTLRYHTGSLAFGALILTLVQIARVIL  
EYIDHKLRGVQNPVARCIMCCFKCCLWCLEKFIKFLNRNAYIMIAIYGKNFCVSAKNAFMLLMRN  
IVRVVVLDDKVTDLLFFGKLLVVGGVGVLSFFFFSGRIPGLGKDFKSPHLNYYWLPIMTSILGAY  
VIASGFFSVFGMCVDTLFLCFLEDLERNNGSLDRPYMSKSLKILGKKNEAPPDNKKRKK

**Important features:**

**Transmembrane domains:**

amino acids 57-80 (type II), 110-126, 215-231, 254-274

**N-glycosylation sites.**

amino acids 16-20, 27-31, 289-293

**Hypothetical YBR002c family proteins.**

amino acids 276-288

**Ammonium transporters proteins.**

amino acids 204-231

**N-myristoylation sites.**

amino acids 60-66, 78-84

**Amidation site.**

amino acids 306-310

**FIGURE 11**

GCCCCGCGCCCGGCGCCGGGCGCCCGAAGCCGGGAGCCACCGCCATGGGGGCCTGCCTGGGAGCC  
TGCTCCCTGCTCAGCTGCGCGTCCTGCCTCTGCGGCTCTGCCCCCTGCATCCTGTGCAGCTGCTG  
CCCCGCCAGCCGCAACTCCACCGTGAGCCGCCTCATCTTCACGTTCTTCCTCTTCCTGGGGGTGC  
TGGTGTCCATCATTATGCTGAGCCCGGGCGTGAGAGTCAGCTCTACAAGCTGCCCTGGGTGTGT  
GAGGAGGGGGCCGGGATCCCCACCGTCCTGCAGGGCCACATCGACTGTGGCTCCCTGCTTGGCTA  
CCGCGCTGTCTACCGCATGTGCTTCGCCACGGCGGCCTTCTTCTTCTTCTTTTACCCTGCTCA  
TGCTCTGCGTGAGCAGCAGCCGGGACCCCCGGGCTGCCATCCAGAATGGGTTTTGGTTCTTTAAG  
TTCCTGATCCTGGTGGGCCTCACCGTGGGTGCCTTCTACATCCCTGACGGCTCCTTCACCAACAT  
CTGGTTCTACTTCGGCGTCGTGGGCTCCTTCCTCTTCATCCTCATCCAGCTGGTGTGCTCATCG  
ACTTTGCGCACTCCTGGAACCAGCGGTGGCTGGGCAAGGCCGAGGAGTGCATTCCTCGTGCCTGG  
TACGCAGGCCTCTTCTTCTTCACTCTCCTCTTCTACTTGCTGTGATCGCGGCCGTGGCGCTGAT  
GTTATGTACTACACTGAGCCAGCGGTGCCACGAGGGCAAGGTCTTCATCAGCCTCAACCTCA  
CCTTCTGTGTCTGCGTGTCATCGCTGCTGTCTGCCCCAAGGTCCAGGACGCCAGCCCAACTCG  
GGTCTGCTGCAGGCCTCGGTCATCACCTCTACACCATGTTTGTACCTGGTCAGCCCTATCCAG  
TATCCCTGAACAGAAATGCAACCCCCATTTGCCAACCAGCTGGGCAACGAGACAGTTGTGGCAG  
GCCCCGAGGGCTATGAGACCCAGTGGTGGGATGCCCCGAGCATTGTGGGCCTCATCATCTTCCTC  
CTGTGCACCCTCTTCATCAGTCTGCGCTCCTCAGACCACCGGCAGGTGAACAGCCTGATGCAGAC  
CGAGGAGTGCCACCTATGCTAGACGCCACACAGCAGCAGCAGCAGCAGGTGGCAGCCTGTGAGG  
GCCGGGCCTTTGACAACGAGCAGGACGGCGTCACCTACAGCTACTCCTTCTTCCACTTCTGCCTG  
GTGCTGGCCTCACTGCACGTATGATGACGCTCACCAACTGGTACAAGCCCGGTGAGACCCGGAA  
GATGATCAGCACGTGGACCGCCGTGTGGGTGAAGATCTGTGCCAGCTGGGCAGGGCTGCTCCTCT  
ACCTGTGGACCCTGGTAGCCCCACTCCTCCTGCGCAACCGCGACTTCAGCTGAGGCAGCCTCACA  
GCCTGCCATCTGGTGCCTCCTGCCACCTGGTGCCTCTCGGCTCGGTGACAGCCAACTGCCCCCT  
CCCCACACCAATCAGCCAGGCTGAGCCCCACCCCTGCCCCAGCTCCAGGACCTGCCCCCTGAGCC  
GGGCCTTCTAGTCGTAGTGCTTCAGGGTCCGAGGAGCATCAGGCTCCTGCAGAGCCCCATCCCC  
CCGCCACACCCACACGGTGGAGCTGCCTCTTCTTCCCCCTCCTCCCTGTTGCCCATACTCAGCAT  
CTCGGATGAAAGGGCTCCCTTGTCCTCAGGCTCCACGGGAGCGGGGCTGCTGGAGAGAGCGGGGA  
ACTCCCACACAGTGGGGCATCCGGCACTGAAGCCCTGGTGTTCCTGGTCACGTCCCCCAGGGGA  
CCCTGCCCCCTTCTGGACTTCGTGCCTTACTGAGTCTCTAAGACTTTTTCTAATAAACAAGCCA  
GTGCGTGTAATAAAAAA

## **FIGURE 12**

MGACLGACSLLSASCCLCGSAPCILCSCCPASRNSTVSRLI FTFFLFLGVLVSIIMLS PGVESQL  
YKLPWVCEEGAGIPTVLQGHIDCGSLLGYRAVYRMC FATAAFFFFFFFFTLLMLCVSSSRDPRAAIQ  
NGFWFFKFLILVGLTVGA FYIPDGSFTNIWFYFGVVGSFLFILIQLVLLIDFAHSWNQRWLGKAE  
ECDSRAWYAGLFFFTLLFYLLSIAAVALMFMYYTEPSGCHEGKVFISLNLTFVCVVSIAAVLPKV  
QDAQPNSGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVAGPEGYETQWW DAPSI  
VGLIIFLLCTLFISLRSSDHRQVNSLMQTEECPPMLDATQQQQQVAACEGRAFDNEQDGV TYSY  
SFFHFCLVLASLHVMMTLTNWYKPGETRKMISTWTAVVWKICASWAGLLLYLWTLVAPLLL RNRD  
FS

**Signal sequence:**

amino acids 1-20

**Transmembrane domains:**

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,  
272-283, 324-340, 391-406, 428-444



**FIGURE 13**

CGGGCCAGCCTGGGGCGGGCCGGCCAGGAACCAACCCGTTAAGGTGTCTTCTCTTTAGGGATGGTGA  
GGTTGGAAAAAGACTCCTGTAACCCTCCTCCAGGATGAACCACCTGCCAGAAGACATGGAGAACG  
CTCTCACCGGGAGCCAGAGCTCCCATGCTTCTCTGCGCAATATCCATTCCATCAACCCACACAA  
CTCATGGCCAGGATTGAGTCCTATGAAGGAAGGGAAAAGAAAGGCATATCTGATGTCAGGAGGAC  
TTTCTGTTTGTGGTGTACCTTTGACCTCTTATTCGTAACATTACTGTGGATAATAGAGTTAAATG  
TGAATGGAGGCATTGAGAACACATTAGAGAAGGAGGTGATGCAGTATGACTACTATTCTTCATAT  
TTTGATATATTTCTTCTGGCAGTTTTTTCGATTTAAAGTGTTAATACTTGCATATGCTGTGTGCAG  
ACTGCGCCATTGGTGGGCAATAGCGTTGACAACGGCAGTGACCAGTGCCTTTTTACTAGCAAAAG  
TGATCCTTTCGAAGCTTTTCTCTCAAGGGGCTTTTGGCTATGTGCTGCCCATCATTTTCATTCATC  
CTTGCCCTGGATTGAGACGTGGTTCCTGGATTTCAAAGTGTTACCTCAAGAAGCAGAAGAAGAAAA  
CAGACTCCTGATAGTTCAGGATGCTTCAGAGAGGGCAGCACTTATACCTGGTGGTCTTTCTGATG  
GTCAGTTTTATTCCCCTCCTGAATCCGAAGCAGGATCTGAAGAAGCTGAAGAAAAACAGGACAGT  
GAGAAACCACTTTTAGAACTATGAGTACTACTTTTGTTAAATGTGAAAAACCCTCACAGAAAGTC  
ATCGAGGCCAAAAGAGGCAGGCAGTGGAGTCTCCCTGTGACAGTAAAGTTGAAATGGTGACGTC  
CACTGCTGGCTTTATTGAACAGCTAATAAAGATTTATTTATTGTAATACCTCACAAACGTTGTAC  
CATATCCATGCACATTTAGTTGCCTGCCTGTGGCTGGTAAGGTAATGTCATGATTCATCCTCTCT  
TCAGTGAGACTGAGCCTGATGTGTTAACAAATAGGTGAAGAAAGTCTTGTGCTGTATTCCTAATC  
AAAAGACTTAATATATTGAAGTAACACTTTTTTAGTAAGCAAGATACCTTTTTTATTTCAATTCAC  
AGAATGGAATTTTTTTGTTTCATGTCTCAGATTTATTTTGTATTTCTTTTTTAACACTCTACATT  
TCCCTTGTTTTTTAACTCATGCACATGTGCTCTTTGTACAGTTTTAAAAAGTGTAATAAAATCTG  
ACATGTCAATGTGGCTAGTTTTATTTTTCTTGTTTTGCATTATGTGTATGGCCTGAAGTGTTGGA  
CTTGCAAAAGGGGAAGAAAGGAATTGCGAATACATGTAAAATGTCACCAGACATTTGTATTATTT  
TTATCATGAAATCATGTTTTCTCTGATTGTTCTGAAATGTTCTAAATACTCTTATTTTGAATGC  
ACAAAATGACTTAAACCATTTCATATCATGTTTCCTTTGCGTTCAGCCAATTTCAATTAAAATGAA  
CTAAATTAAAAA

## **FIGURE 14**

MNHLPEDMENALTGSQSSHASLRNIHSINPTQLMARIESYEGREKKGISDVRRTFCLFVTFDLLF  
VTLLWIIELNVNGGIENLEKEVMQYDYYSSYFDIFLLAVFRFKVLILAYAVCRLRHWWAIALTT  
AVTSAFLLAKVILSKLFSQGAFGYVLPPIISFILAWIETWFLDFKVLQPQAEENRLLIVQDASER  
AALIPGGLSDGQFYSPPESEAGSEEAEEKQDSEKPLLEL

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-20

#### **Transmembrane domains:**

amino acids 54-72, 100-118, 130-144, 146-166

#### **N-myristoylation sites.**

amino acids 14-20, 78-84, 79-85, 202-208, 217-223

**FIGURE 15**

ACTCGAACGCAGTTGCTTCGGGACCCAGGACCCCCTCGGGCCCCGACCCGCCAGGAAAGACTGAGG  
 CCGCGGCCTGCCCCGCCCCGGCTCCCTGCGCCGCGCCGCCTCCCGGGACAGAAG**ATGT**GCTCCAG  
 GGTCCCTCTGCTGCTGCCGCTGCTCCTGCTACTGGCCCTGGGGCCTGGGGTGCAGGGCTGCCCAT  
 CCGGCTGCCAGTGCAGCCAGCCACAGACAGTCTTCTGCACTGCCCGCCAGGGGACCACGGTGCCC  
 CGAGACGTGCCACCCGACACGGTGGGGCTGTACGTCTTTGAGAACGGCATCACCATGCTCGACGC  
 AGGCAGCTTTGCCGGCCTGCCGGGCCTGCAGCTCCTGGACCTGTACAGAACCCAGATCGCCAGCC  
 TGCCAGCGGGGTCTTCCAGCCACTCGCCAACCTCAGCAACCTGGACCTGACGGCCAACAGGCTG  
 CATGAAATCACCAATGAGACCTTCCGTGGCCTGCGGCGCCTCGAGCGCCTCTACCTGGGCAAGAA  
 CCGCATCCGCCACATCCAGCCTGGTGCCTTCGACACGCTCGACCGCCTCCTGGAGCTCAAGCTGC  
 AGGACAACGAGCTGCGGGCACTGCCCCCGCTGCGCCTGCCCGCCTGCTGCTGCTGGACCTCAGC  
 CACAACAGCCTCCTGGCCCTGGAGCCCGGCATCCTGGACACTGCCAACGTGGAGGCGCTGCGGCT  
 GGCTGGTCTGGGGCTGCAGCAGCTGGACGAGGGGCTCTTCAGCCGCTTGCGCAACCTCCACGACC  
 TGGATGTGTCCGACAACCAGCTGGAGCGAGTGCCACCTGTGATCCGAGGCCTCCGGGGCCTGACG  
 CGCCTGCGGCTGGCCGGCAACACCCGCATTGCCAGCTGCGGCCCGAGGACCTGGCCGGCCTGGC  
 TGCCCTGCAGGAGCTGGATGTGAGCAACCTAAGCCTGCAGGCCCTGCCTGGCGACCTCTCGGGCC  
 TCTTCCCCCGCCTGCGGCTGCTGGCAGCTGCCCGCAACCCCTTCAACTGCGTGTGCCCCCTGAGC  
 TGGTTTGGCCCCCTGGGTGCGCGAGAGCCACGTCACTGGCCAGCCCTGAGGAGACGCGCTGCCA  
 CTTCCCGCCCAAGAACGCTGGCCGGCTGCTCCTGGAGCTTGACTACGCCGACTTTGGCTGCCCAG  
 CCACCACCACACAGCCACAGTGCCACACAGAGGCCCGTGGTGCGGGAGCCCACAGCCTTGTCT  
 TCTAGCTTGGCTCCTACCTGGCTTAGCCCCACAGCGCCGGCCACTGAGGCCCCAGCCCGCCCTC  
 CACTGCCCCACCGACTGTAGGGCCTGTCCCCAGCCCCAGGACTGCCACCGTCCACCTGCCTCA  
 ATGGGGGCACATGCCACCTGGGGACACGGCACCCACCTGGCGTGCTTGTGCCCCGAAGGCTTCACG  
 GGCCTGTACTGTGAGAGCCAGATGGGGCAGGGGACACGGCCAGCCCTACACCAGTCACGCCGAG  
 GCCACCACGGTCCCTGACCCTGGGCATCGAGCCGGTGAGCCCCACCTCCCTGCGCGTGGGGCTGC  
 AGCGCTACCTCCAGGGGAGCTCCGTGCAGCTCAGGAGCCTCCGTCTCACCTATCGCAACCTATCG  
 GGCCCTGATAAGCGGCTGGTGACGCTGCGACTGCCTGCCTCGCTCGCTGAGTACACGGTCACCCA  
 GCTGCGGGCCCAACGCCACTTACTCCGTCTGTGTATGCCTTTGGGGCCCGGGCGGGTGCCGGAGG  
 GCGAGGAGGCCTGCGGGGAGGCCATACACCCCCAGCCGTCCACTCCAACCACGCCCCAGTCACC  
 CAGGCCCGCGAGGGCAACCTGCCGCTCCTCATTGCGCCCGCCCTGGCCGCGGTGCTCCTGGCCGC  
 GCTGGCTGCGGTGGGGGCAGCCTACTGTGTGCGGCGGGGGCGGGCCATGGCAGCAGCGGCTCAGG  
 ACAAAGGGCAGGTGGGGCCAGGGGCTGGGCCCCTGGAACGGAGGGAGTGAAGGTCCCCTTGAGAG  
 CCAGGCCCGAAGGCAACAGAGGGCGGTGGAGAGGCCCTGCCAGCGGGTCTGAGTGTGAGGTGCC  
 ACTCATGGGCTTCCCAGGGCCTGGCCTCCAGTCACCCCTCCACGCAAAGCCCTACATC**TAAG**CCA  
 GAGAGAGACAGGGCAGCTGGGGCCGGGCTCTCAGCCAGTGAGATGGCCAGCCCCCTCCTGCTGCC  
 ACACCACGTAAGTTCTCAGTCCCAACCTCGGGGATGTGTGCAGACAGGGCTGTGTGACCACAGCT  
 GGGCCCTGTTCCCTCTGGACCTCGGTCTCCTCATCTGTGAGATGCTGTGGCCCAGCTGACGAGCC  
 CTAACGTCCCCAGAACCGAGTGCCATATGAGGACAGTGTCCGCCCTGCCCTCCGCAACGTGCAGTC  
 CCTGGGCACGGCGGGGCCCTGCCATGTGCTGGTAACGCATGCCTGGGTCTGCTGGGCTCTCCAC  
 TCCAGGCGGACCCTGGGGGCCAGTGAAGGAAGCTCCCGGAAAGAGCAGAGGGAGAGCGGGTAGGC  
 GGCTGTGTGACTCTAGTCTTGGCCCCAGGAAGCGAAGGAACAAAGAAACTGGAAAGGAAGATGC  
 TTTAGGAACATGTTTTGCTTTTTTAAATATATATATTTATAAGAGATCCTTTCCCATTTATTCT  
 GGAAGATGTTTTTCAAACCTCAGAGACAAGGACTTTGGTTTTTGTAAAGACAAACGATGATATGAA  
 GGCCTTTTGTAAAGAAAAAATAAAAGATGAAGTGTGAAA

## **FIGURE 16**

MCSRVPLLLPLLLLLALGPGVQGCPSGCQCSQPQTVFCTARQGTTVPRDVPPDTVGLYVFENGIT  
 MLDAGSFAGLPGLQLLDLSQNQIASLPSGVFQPLANLSNLDLTANRLHEITNETFRGLRRRLERLY  
 LGKNRIRHIQPGAFDTLDRLLELKLQDNELRALPPLRLPRLLLDLSHNSLLALEPGILDTANVE  
 ALRLAGLGLQQLDEGLFSRLRNLDLSDNQLERVPPVIRGLRGLTRLRLAGNTRIAQLRPEDL  
 AGLAALQELDVSNSLSLQALPGDLSGLFPRLRLLAAARNPFNCVCPLSWFGPWVRESHVTLASPEE  
 TRCHFPPKNAGRLLLELDYADFGCPATTTTATVPTTRPVVREPTALSSSLAPTWLSPTAPATEAP  
 SPPSTAPPTVGVPVQPQDCPPSTCLNGGTCHLGTRHHLACLCPEGFTGLYCESQMGQGTRPSPTP  
 VTPRPPRSLTLGIEPVSPSTLRVGLQRYLQGSSVQLRSLRLTYRNLSGPDKRLVTLRLPASLAEY  
 TVTQLRPNATYSVCMPLGPGRVPEGEEACGEAHTPPAVHSNHAPVTQAREGNLPLLIAPALAAV  
 LLAALAAVGAAYCVRGRAMAAAAQDKGQVGPGAGPLELEGVKVPLEPGPKATEGGGEALPSGSE  
 CEVPLMGFPGPGLQSPLHAKPYI

### **Important features:**

#### **Signal peptide:**

amino acids 1-23

#### **Transmembrane domain:**

amino acids 579-599

#### **EGF-like domain cysteine pattern signature.**

amino acids 430-442

#### **Leucine zipper pattern.**

amino acids 197-219, 269-291

#### **N-glycosylation sites.**

amino acids 101-105, 117-121, 273-277, 500-504, 528-532

#### **Tyrosine kinase phosphorylation sites.**

amino acids 124-131, 337-345

#### **N-myristoylation sites.**

amino acids 23-29, 27-33, 70-76, 142-148, 187-193, 348-354, 594-600,  
 640-646

[illegible]

## **FIGURE 18**

MRVRIGLTLLLCVLLSLASASSDEEGSQDESLDSKTTLTSDSVKDHTTAGRVVAGQIFLDSEE  
SELESSIQEEEDSLKSQEGESVTEDISFLESPNPENKDYEEPKKVRKPALTAIEGTAHGEPCHFP  
FLFLDKEYDECTSDGREDGRLWCATTYDYKADEKWGFCETEEEEAAKRRQMQEAMMYQTGMKILN  
GSNKKSQKREAYRYLQKAASMNHTKALERVSYALLFGDYLPQNIQAAREMFEEKLTEEGSPKGQTA  
LGFLYASGLGVNSSQAKALVYYTFGALGGNLI AHMVLVSRL

**Important features:**

**Signal peptide:**

amino acids 1-21

**N-glycosylation sites.**

amino acids 195-199, 217-221, 272-276

**Tyrosine kinase phosphorylation site.**

amino acids 220-228

**N-myristoylation sites.**

amino acids 120-126, 253-259, 268-274, 270-274, 285-291, 289-295

**Glycosaminoglycan attachment site.**

amino acids 267-271

**Microbodies C-terminal targeting signal.**

amino acids 299-303

**Type II fibronectin collagen-binding domain protein.**

amino acids 127-169

**Fructose-bisphosphate aldolase class-II protein.**

amino acids 101-119

**FIGURE 19**

AATTCAGATTTTAAAGCCCATTCCTGCAGTGGAATTTTCATGAACTAGCAAGAGGACACCATCTTCTT  
GTATTATACAAGAAAGGAGTGTACCTATCACACACAGGGGGAAAA**ATG**CTCTTTTGGGTGCTAGG  
CCTCCTAATCCTCTGTGGTTTTCTGTGGACTCGTAAAGGAAAACTAAAGATTGAAGACATCACTG  
ATAAGTACATTTTTATCACTGGATGTGACTCGGGCTTTGGAACTTGGCAGCCAGAAGCTTTTGAT  
AAAAAGGGATTTTCATGTAATCGCTGCCTGTCTGACTGAATCAGGATCAACAGCTTTAAAGGCAGA  
AACCTCAGAGAGACTTCGTACTGTGCTTCTGGATGTGACCGACCCAGAGAATGTCAAGAGGACTG  
CCCAGTGGGTGAAGAACCAAGTTGGGGAGAAAGGTCTCTGGGGTCTGATCAATAATGCTGGTGTT  
CCCGGCGTGCTGGCTCCCACTGACTGGCTGACACTAGAGGACTACAGAGAACCTATTGAAGTGAA  
CCTGTTTGGACTCATCAGTGTGACACTAAATATGCTTCCTTTGGTCAAGAAAGCTCAAGGGAGAG  
TTATTAATGTCTCCAGTGTTGGAGGTCGCCTTGCAATCGTTGGAGGGGGCTATACTCCATCCAAA  
TATGCAGTGGAAGGTTTCAATGACAGCTTAAGACGGGACATGAAAGCTTTTGGTGTGCACGTCTC  
ATGCATTGAACCAGGATTGTTCAAAACAAACTTGGCAGATCCAGTAAAGGTAATTGAAAAAAAC  
TCGCCATTTGGGAGCAGCTGTCTCCAGACATCAAACAACAATATGGAGAAGGTTACATTGAAAAA  
AGTCTAGACAACTGAAAGGCAATAAATCCTATGTGAACATGGACCTCTCTCCGGTGGTAGAGTG  
CATGGACCACGCTCTAACAAGTCTCTTCCCTAAGACTCATTATGCCGCTGGAAAAGATGCCAAAA  
TTTTCTGGATACCTCTGTCTCACATGCCAGCAGCTTTGCAAGACTTTTTATTGTTGAAACAGAAA  
GCAGAGCTGGCTAATCCCAAGGCAGT**TGA**CTCAGCTAACCACAAATGTCTCCTCCAGGCTATGA  
AATTGGCCGATTTCAAGAACACATCTCCTTTTCAACCCCATTCCTTATCTGCTCCAACCTGGACT  
CATTTAGATCGTGCTTATTTGGATTGCAAAGGGAGTCCCACCATCGCTGGTGGTATCCCAGGGT  
CCCTGCTCAAGTTTTCTTTGAAAAGGAGGGCTGGAATGGTACATCACATAGGCAAGTCCTGCCCT  
GTATTTAGGCTTTGCCTGCTTGGTGTGATGTAAGGGAAATTGAAAGACTTGCCCATTCAAAATGA  
TCTTTACCGTGGCCTGCCCCATGCTTATGGTCCCCAGCATTTACAGTAACTTGTGAATGTTAAGT  
ATCATCTCTTATCTAAATATTAAAAGATAAGTCAACCCAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAA

## **FIGURE 20**

MLFWVLGLLILCGFLWTRKGKLIKIEDITDKYIFITGCDSGFGNLAARTFDKKGFHVIAACLTESG  
STALKAETSERLRTVLLDVTDPENVKRTAQWVKNQVGEKGLWGLINNAGVPGVLAPTDWLTLEDY  
REPIEVNLFGLISVTLNMLPLVKKAQGRVINVSSVGGRLAIVGGGYTPSKYAVEGFNDLRRDMK  
AFGVHVSCIEPGLFKTNLADPVKVIEKKLAIWEQLSPDIKQQYGEGYIEKSLDKLKGNKSYVNMD  
LSPVVECMDHALTSLFPKTHYAAGKDAKIFWIPLSHMPAALQDFLLLKQKAELANPKAV

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-17

#### **Transmembrane domain:**

amino acids 136-152

#### **N-glycosylation sites.**

amino acids 161-163, 187-190 and 253-256

#### **Glycosaminoglycan attachment site.**

amino acids 39-42

#### **N-myristoylation sites.**

amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212



**FIGURE 21**

CTGAGGCGGCGGTAGC**ATG**GAGGGGGAGAGTACGTCGGCGGTGCTCTCGGGCTTTGTGCTCGGCG  
CACTCGCTTTCCAGCACCTCAACACGGACTCGGACACGGAAGGTTTTCTTCTTGGGGAAGTAAAA  
GGTGAAGCCAAGAACAGCATTACTGATTCCCAAATGGATGATGTTGAAGTTGTTTATACAATTGA  
CATTCAGAAATATATTCCATGCTATCAGCTTTTTAGCTTTTATAATTCTTCAGGCGAAGTAAATG  
AGCAAGCACTGAAGAAAATATTATCAAATGTCAAAAAGAATGTGGTAGGTTGGTACAAATTCCGT  
CGTCATTTCAGATCAGATCATGACGTTTAGAGAGAGGCTGCTTCACAAAACTTGCAGGAGCATTT  
TTCAAACCAAGACCTTGTTTTTCTGCTATTAACACCAAGTATAATAACAGAAAGCTGCTCTACTC  
ATCGACTGGAACATTCCTTATATAAACCTCAAAAAGGACTTTTTTCACAGGGTACCTTTAGTGTT  
GCCAATCTGGGCATGTCTGAACAACTGGGTTATAAACTGTATCAGGTTCTGTATGTCCACTGG  
TTTTAGCCGAGCAGTACAAACACACAGCTCTAAATTTTTTGAAGAAGATGGATCCTTAAAGGAGG  
TACATAAGATAAATGAAATGTATGCTTCATTACAAGAGGAATTAAAGAGTATATGCAAAAAAGTG  
GAAGACAGTGAACAAGCAGTAGATAAACTAGTAAAGGATGTAAACAGATTAAAACGAGAAATTGA  
GAAAAGGAGAGGAGCACAGATTCAGGCAGCAAGAGAGAAGAACATCCAAAAGACCCTCAGGAGA  
ACATTTTTCTTTGTCAGGCATTACGGACCTTTTTTCCAAATTCTGAATTTCTTCATTCATGTGTT  
ATGTCTTTAAAAAATAGACATGTTTCTAAAAGTAGCTGTAAC TACAACCACCATCTCGATGTAGT  
AGACAATCTGACCTTAATGGTAGAACACACTGACATTCCTGAAGCTAGTCCAGCTAGTACACCAC  
AAATCATTAAGCATAAAGCCTTAGACTTAGATGACAGATGGCAATTCAAGAGATCTCGGTTGTTA  
GATACACAAGACAAACGATCTAAAGCAAATACTGGTAGTAGTAACCAAGATAAAGCATCCAAAAT  
GAGCAGCCCAGAAACAGATGAAGAAATTGAAAAGATGAAGGGTTTTTGGTGAATATTCACGGTCTC  
CTACATTT**TGA**TCCTTTTAAACCTTACAAGGAGATTTTTTTATTGGCTGATGGGTAAAGCCAAAC  
ATTTCTATTGTTTTTACTATGTTGAGCTACTTGCAGTAAGTTCATTTGTTTTTACTATGTTCCACC  
TGTTTGCAGTAATACACAGATAACTCTTAGTGCAATTTACTTCACAAAGTACTTTTTTCAAACATCA  
GATGCTTTTATTTCCAAACCTTTTTTTCACCTTTCACTAAGTTGTTGAGGGGAAGGCTTACACAG  
ACACATTCTTTAGAATTGGAAAAGTGAGACCAGGCACAGTGGCTCACACCTGTAATCCCAGCACT  
TAGGGAAGACAAGTCAGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAGCCTGGGCAACGTATT  
GAGACCATGTCTATTAAAAAATAAAATGGAAAAGCAAGAATAGCCTTATTTTTCAAATATGGAAA  
GAAATTTATATGAAAATTTATCTGAGTCATTAAAATTCTCCTTAAGTGATACTTTTTTAGAAGTA  
CATTATGGCTAGAGTTGCCAGATAAAATGCTGGATATCATGCAATAAATTTGCAAAACATCATCT  
AAAATTTAAAAA

## **FIGURE 22**

MEGESTSAVLSGFVLGALAFQHLNTDSDTEGFLLGEVKGEAKNSITDSQMDDVEVVYTIDIQKYI  
PCYQLFSFYNSSGEVNEQALKKILSNVKNVVGWYKFRRHSDQIMTFRERLLHKNLQEHFSNODL  
VFLLLTPSIITESCSTHRLEHSLYKPQKGLFHRVPLVVANLGMSEQLGYKTVSGSCMSTGFSRAV  
QTHSSKFFFEEDGSLKEVHKINEMYASLQEELKSICKKVEDSEQAVDKLVKDVNRLKREIEKRRGA  
QIQAAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHVSKSSCNYNHHLDVVDNLTL  
MVEHTDIPEASPASTPQIIKHKALDLDDRWQFKRSRLLDLTQDKRSKANTGSSNQDKASKMSSPET  
DEEIEKMKGFGEYSRSPTE

**Important features:**

**Signal peptide:**

amino acids 1-19

**N-glycosylation sites.**

amino acids 75-79, 322-326

**N-myristoylation site.**

amino acids 184-154

**Growth factor and cytokines receptors family.**

amino acids 134-150

**FIGURE 23**

GGCACAGCCGCGCGGGCGGAGGGCAGAGTCAGCCGAGCCGAGTCCAGCCGGACGAGCGGACCAGCGCAGGGCAGCCCAA  
GCAGCGCGCAGCGAACGCCCCGCCGCCGCCACACCCTCTGCGGTCCCCGCGGCGCCTGCCACCCTTCCCTCCTTCCCC  
GCGTCCCCGCTCGCCGGCCAGTCAGCTTGCCGGGTTCGCTGCCCCGCGAAACCCGAGGTACCAGCCCGCGCCTCT  
GCTTCCCTGGGCGCGCGCCGCTCCACGCCCTCCTTCTCCCTGGCCCGGCGCCTGGCACCAGGGGACCGTTGCCTGA  
CGCGAGGCCCAGCTCTACTTTTCGCCCCGCTCTCCTCCGCCTGCTCGCCTCTTCCACCAACTCCAACCTCCTTCTCCC  
TCCAGCTCCACTCGCTAGTCCCCGACTCCGCCAGCCCTCGGCCCGCTGCCGTAGCGCCGCTTCCCGTCCGGTCCCAAA  
GGTGGGAACGCGTCCGCCCGGCCCGCACCATGGCACGGTTCGGCTTGCCCGCGCTTCTCTGCACCCTGGCAGTGCTC  
AGCGCCGCGCTGCTGGCTGCCGAGCTCAAGTCGAAAAGTTGCTCGGAAGTGCGACGTCTTTACGTGTCCAAAGGCTTC  
AACAAGAACGATGCCCCCTCCACGAGATCAACGGTGATCATTTGAAGATCTGTCCCCAGGGTTCTACCTGCTGCTCT  
CAAGAGATGGAGGAGAAGTACAGCCTGCAAAGTAAAGATGATTTCAAAGTGTGGTCAGCGAACAGTGCAATCATTTG  
CAAGCTGTCTTTGCTTACGTTACAAGAAGTTTGATGAATTCTTCAAAGAACTACTTGAAAATGCAGAGAAATCCCTG  
AATGATATGTTTGTGAAGACATATGGCCATTTATACATGCAAAATCTTGAGCTATTTAAAGATCTCTTCGTAGAGTTG  
AAACGTTACTACGTGGTGGGAAATGTGAACCTGGAAGAAATGCTAAATGACTTCTGGGCTCGCCTCCTGGAGCGGATG  
TTCCGCCTGGTGAACCTCCAGTACCACTTTACAGATGAGTATCTGGAATGTGTGAGCAAGTATACGGAGCAGCTGAAG  
CCCTTCGGAGATGTCCCTCGCAAATTGAAGCTCCAGGTTACTCGTGCTTTTGTAGCAGCCCGTACTTTTCGCTCAAGGC  
TTAGCGGTTGCGGGAGATGTCTGTGAGCAAGGTCTCCGTGGTAAACCCACAGCCCAGTGATCCCATGCCCTGTTGAAG  
ATGATCTACTGCTCCCACTGCCGGGGTCTCGTGACTGTGAAGCCATGTTACAACACTACTGCTCAAACATCATGAGAGGC  
TGTTTTGGCCAACCAAGGGGATCTCGATTTTGAATGGAACAATTTTCATAGATGCTATGCTGATGGTGGCAGAGAGGCTA  
GAGGGTCCTTTCAACATTGAATCGGTTCATGGATCCCATCGATGTGAAGATTTCTGATGCTATTATGAACATGCAGGAT  
AATAGTGTTCAAGTGTCTCAGAAGGTTTTCCAGGGATGTGGACCCCCAAGCCCCCTCCAGCTGGACGAATTTCTCGT  
TCCATCTCTGAAAGTGCCTTCAGTGCTCGCTTCAGACCACATACCCCCGAGGAACGCCCAACCACAGCAGCTGGCACT  
AGTTTGGACCGACTGGTTACTGATGTCAAGGAGAACTGAAACAGGCCAAGAAATTTCTGGTCTCCCTTCCGAGCAAC  
GTTTGAACGATGAGAGGATGGCTGCAGGAAACGGCAATGAGGATGACTGTTGGAATGGGAAAGGCAAAAGCAGGTAC  
CTGTTTGCAGTGACAGGAAATGGATTAGCCAACCAGGGCAACAACCCAGAGGTCCAGGTTGACACCAGCAAACCAGAC  
ATACTGATCCTTCGTCAAATCATGGCTCTTCGAGTGATGACCAGCAAGATGAAGAATGCATACAATGGGAACGACGTG  
GACTTCTTTGATATCAGTGATGAAAGTAGTGAGAGGAAGTGGAAAGTGGCTGTGAGTATCAGCAGTGCCCTTCAGAG  
TTTGACTACAATGCCACTGACCATGCTGGGAAGAGTGCCAATGAGAAAGCCGACAGTGCTGGTGTCCGTCTGGGGCA  
CAGGCCTACCTCCTCACTGTCTTCTGCATCTTGTTCTGGTTATGCAGAGAGAGTGGAGATTAATTCTCAAACCTCTGAG  
AAAAAGTGTTTCATCAAAAAGTTAAAAGGCACCAGTTATCACTTTTCTACCATCCTAGTGACTTTTGCTTTTTAAATGAA  
TGGACAACAATGTACAGTTTTTACTATGTGGCCACTGGTTTAAAGAGTGCTGACTTTGTTTTCTCATTAGTTTTGGG  
AGGAAAAGGGACTGTGCATTGAGTTGGTTCTGCTCCCCCAAACCATGTTAAACGTGGCTAACAGTGTAGGTACAGAA  
CTATAGTTAGTTGTGCATTTGTGATTTTATCACTCTATTATTTGTTTGTATGTTTTTTCTCATTTCTGTTTGTGGGTT  
TTTTTTTCCAACGTGTGATCTCGCCTTGTTTCTTACAAGCAAACCAGGGTCCCTTCTTGGCACGTAACATGTACGTATT  
TCTGAAATATTAAATAGCTGTACAGAAGCAGGTTTTATTTATCATGTTATCTTATTTAAAGAAAAAGCCCAAAAGC

## **FIGURE 24**

MARFGLPALLCTLAVLSAALLAAELKSKSCSEVRRRLYVSKGFNKNDAPLHEINGDHLKICPQGST  
CCSQEMEEEKYSLQSKDDFKSVVSEQCNHLQAVFASRYKKFDEFFKELLENAEKSLNDMFVKTYGH  
LYMQNSELFKDLFVELKRYVVGNNLEEMLNDFWARLLERMFLVNSQYHFTDEYLECVSKYTE  
QLKPFQDVPRKLKLQVTRAFVAARTFAQGLAVAGDVVSKVSNNPTAQCTHALLKMIYCSHCRGL  
VTVKPCYNYCSNIMRGCLANQGDLDFFWNNFIDAMLMVAERLEGPFNIESVMDPIDVKISDAIMN  
MQDNSVQVSQKVFQCGPPKPLPAGRISRSISESAFSARFRPHHPEERPTTAAGTSLDRLVTDVK  
EKLKQAKKFWSSLPSNVCNDERMAAGNGNEDDCWNGKGSRYLFAVTGNGLANQGNNPEVQVDT  
KPDILILRQIMALRVMTSKMKNAYNGNDVDFDISDESSGEGSGSGCEYQQCPSEFDYNATDHAG  
KSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

**Important features:**

**Signal peptide:**

amino acids 1-22

**ATP/GTP-binding site motif A (P-loop).**

amino acids 515-524

**N-glycosylation site.**

amino acids 514-518

**Glycosaminoglycan attachment sites.**

amino acids 494-498, 498-502

**N-myristoylation sites.**

amino acids 63-69, 224-230, 276-282, 438-444, 497-503, 531-537

**Glypicans proteins.**

amino acids 54-75, 105-157, 238-280, 309-346, 423-460, 468-506

**FIGURE 25**

CTCGCCCTCAAATGGGAACGCTGGCCTGGGACTAAAGCATAGACCACCAGGCTGAGTATCCTGAC  
CTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTTCCATTATATTCTTCAAGCAACT  
TACAGCTGCACCGACAGTTGCGATGAAAGTTCTAATCTCTTCCCTCCTCCTGTTGCTGCCACTAA  
TGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGGTCGCCAGAGGCCACAGGGACCGAGGC  
CAGGCTTCTAGGAGATGGCTCCAGGAAGGCGGCCAAGAATGTGAGTGCAAAGATTGGTTCCTGAG  
AGCCCCGAGAAGAAAATTCATGACAGTGTCTGGGCTGCCAAAGAAGCAGTGCCCCTGTGATCATT  
TCAAGGGCAATGTGAAGAAAACAAGACACCAAAGGCACCACAGAAAGCCAAACAAGCATTCCAGA  
GCCTGCCAGCAATTTCTCAAACAATGTCAGCTAAGAAGCTTTGCTCTGCCTTTGTAGGAGCTCTG  
AGCGCCCACTCTTCCAATTAAACATTCTCAGCCAAGAAGACAGTGAGCACACCTACCAGACACTC  
TTCTTCTCCACCTCACTCTCCCACTGTACCCACCCCTAAATCATTCCAGTGCTCTCAAAAAGCA  
TGTTTTTCAAGATCATTTTGTTTGTTGCTCTCTCTAGTGTCTTCTCTCTCGTCAGTCTTAGCCT  
GTGCCCTCCCCTTACCCAGGCTTAGGCTTAATTACCTGAAAGATTCCAGGAACTGTAGCTTCCT  
AGCTAGTGTCATTTAACCTTAAATGCAATCAGGAAAGTAGCAAACAGAAGTCAATAAATATTTTT  
AAATGTCAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 26**

MKVLISLLLLLLPLMLMSMVSSSLNPGVARGHRDRGQASRRWLQEGGQECECKDWFLRAPRRKFM  
TVSGLPKKQCPDHFKGNVKKTRHQRHHRKPNKHSRACQQFLKQCQLRSFALPL

**Important features:**

**Signal peptide:**

amino acids 1-22

**N-myristoylation sites.**

amino acids 27-33, 46-52

**FIGURE 27**

GGACGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTGCCCCAGCGGAAGCACAGCTCAG  
AGCTGGTCTGCC**ATG**GACATCCTGGTCCCCTCCTGCAGCTGCTGGTGCTGCTTCTTACCCTGCC  
CCTGCACCTCATGGCTCTGCTGGGCTGCTGGCAGCCCCCTGTGCAAAGCTACTTCCCCTACCTGA  
TGGCCGTGCTGACTCCCAAGAGCAACCGCAAGATGGAGAGCAAGAAACGGGAGCTCTTCAGCCAG  
ATAAAGGGGCTTACAGGAGCCTCCGGGAAAGTGGCCCTACTGGAGCTGGGCTGCGGAACCGGAGC  
CAACTTTCAGTTCTACCCACCGGGCTGCAGGGTCACCTGCCTAGACCCAAATCCCCACTTTGAGA  
AGTTCCTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATATGAGCGGTTTGTGGTGGCTCCT  
GGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGTGGTGGTCTGCACTCTGGTGCTGTG  
CTCTGTGCAGAGCCCAAGGAAGTCCCTGCAGGAGGTCCGGAGAGTACTGAGACCGGGAGGTGTGC  
TCTTTTTCTGGGAGCATGTGGCAGAACCATATGGAAGCTGGGCCTTCATGTGGCAGCAAGTTTTCT  
GAGCCACCTGGAAACACATTGGGGATGGCTGCTGCCTCACCAGAGAGACCTGGAAGGATCTTGA  
GAACGCCCAGTTCTCCGAAATCCAAATGGAACGACAGCCCCCTCCCTTGAAGTGGCTACCTGTTG  
GGCCCCACATCATGGGAAAGGCTGTCAAACAATCTTTCCCAAGCTCCAAGGCACTCATTTGCTCC  
TTCCCCAGCCTCCAATTAGAACAAGCCACCCACCAGCCTATCTATCTTCCACTGAGAGGGACCT**TA**  
**G**CAGAATGAGAGAAGACATTCATGTACCACCTACTAGTCCCTCTCTCCCCAACCTCTGCCAGGGC  
AATCTCTAACTTCAATCCCGCCTTCGACAGTGAAAAAGCTCTACTTCTACGCTGACCCAGGGAGG  
AAACACTAGGACCCTGTTGTATCCTCAACTGCAAGTTTCTGGACTAGTCTCCCAACGTTTGCCTC  
CCAATGTTGTCCCTTTCCCTTCGTTCCCATGGTAAAGCTCCTCTCGCTTTCCTCCTGAGGCTACAC  
CCATGCGTCTCTAGGAACTGGTCACAAAAGTCATGGTGCCTGCATCCCTGCCAAGCCCCCTGAC  
CCTCTCTCCCCACTACCACCTTCTTCCTGAGCTGGGGGCACCAGGGAGAATCAGAGATGCTGGGG  
ATGCCAGAGCAAGACTCAAAGAGGCAGAGGTTTTGTTCTCAAATATTTTTTAATAAATAGACGAA  
ACCACG

## **FIGURE 28**

MDILVPLLQLLVLLLTLPPLHLMALLGCWQPLCKSYFPYLMAVLTPKSNRKMESKKRELFSSQIKGL  
TGASGKVALLELGCGTGANFQFYPPGCRVTCCLDPNPHFEKFLTKSMAENRHLQYERFVVAPGEDM  
RQLADGSMDEVVCTLVLCVQSPRKVLQEVRRVLRPGGVLEFFWEHVAEPYGSWAFMWQQVFEPW  
KHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKWLPGPHIMGKAVKQSFSSKALICSFPSL  
QLEQATHQPIYLPRLGT

**Important features:**

**Signal peptide:**

amino acids 1-23

**Leucine zipper pattern.**

amino acids 10-32

**N-myristoylation sites.**

amino acids 64-70, 78-84, 80-86, 91-97, 201-207



**FIGURE 29**

CAATGTTTGCCTATCCACCTCCCCAAGCCCCTTTACCT**ATG**CTGCTGCTAACGCTGCTGCTGCT  
GCTGCTGCTGCTTAAAGGCTCATGCTTGGAGTGGGGACTGGTCGGTGCCAGAAAGTCTCTTCTG  
CCACTGACGCCCCCATCAGGGATTGGGCCTTCTTTCCCCCTTCCTTTCTGTGTCTCCTGCCTCAT  
CGGCCTGCCATGACCTGCAGCCAAGCCCAGCCCCGTGGGGAAGGGGAGAAAGTGGGGGATGGC**TA**  
**A**GAAAGCTGGGAGATAGGGAACAGAAGAGGGTAGTGGGTGGGCTAGGGGGGCTGCCTTATTTAAA  
GTGGTTGTTTATGATTCTTATACTAATTTATACAAAGATATTAAGGCCCTGTTCAATTAAGAAATT  
GTTCCCTTCCCCTGTGTTCAATGTTTGTAAGATTGTTCTGTGTAAATATGTCTTTATAATAAAC  
AGTTAAAAGCTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 30**

MLLLTLLLLLLLLLKGSCLEWGLVGAQKVSSATDAPIRDWAFPPSFLCLLPHRPAMTCSQAQPRG  
EGEKVGDG

**Important features:**

**Signal peptide:**

amino acids 1-15

**Growth factor and cytokines receptors family:**

amino acids 3-18

**FIGURE 31**

GT TTGAATTCCTTCAACTATAACCCACAGTCCAAAAGCAGACTCACTGTGTCCCAGGCTACCAAGTT  
CCTCCAAGCAAGTCATTTCCCTTATTTAACCGATGTGTCCCTCAAACACCTGAGTGCTACTCCCT  
ATTTGCATCTGTTTTGATAAATGATGTTGACACCCTCCACCGAATTCTAAGTGGAATC**ATG**TCGG  
GAAGAGATACAATCCTTGGCCTGTGTATCCTCGCATTAGCCTTGTCTTTGGCCATGATGTTTACC  
TTCAGATTCATCACCACCCTTCTGGTTCACATTTTCATTTCAATTGGTTATTTTGGGATTGTTGTT  
TGTCTGCGGTGTTTTATGGTGGCTGTATTATGACTATAACCAACGACCTCAGCATAGAATTGGACA  
CAGAAAGGGAAAAATATGAAGTGCGTGCTGGGGTTTGCTATCGTATCCACAGGCATCACGGCAGTG  
CTGCTCGTCTTGATTTTTGTTCTCAGAAAGAGAATAAAATTGACAGTTGAGCTTTTCCAAATCAC  
AAATAAAGCCATCAGCAGTGCTCCCTTCCTGCTGTTCCAGCCACTGTGGACATTTGCCATCCTCA  
TTTTCTTCTGGGTCCTCTGGGTGGCTGTGCTGCTGAGCCTGGGAAGTGCAGGAGCTGCCCAGGTT  
ATGGAAGGCGGCCAAGTGGAATATAAGCCCCCTTCGGGCATTTCGGTACATGTGGTTCGTACCATTT  
AATTGGCCTCATCTGGACTAGTGAATTCATCCTTGCGTGCCAGCAAATGACTATAGCTGGGGCAG  
TGGTTACTTGTTATTTCAACAGAAGTAAAAATGATCCTCCTGATCATCCCATCCTTTTCGTCTCTC  
TCCATTCTCTTCTTCTACCATCAAGGAACCGTTGTGAAAGGGTCATTTTTAATCTCTGTGGTGAG  
GATTCCGAGAATCATTGTCATGTACATGCAAAACGCACTGAAAGAACAGCAGCATGGTGCATTGT  
CCAGGTACCTGTTCCGATGCTGCTACTGCTGTTTCTGGTGTCTTGACAAATACCTGCTCCATCTC  
AACCAGAATGCATATACTACAACCTGCTATTAATGGGACAGATTTCTGTACATCAGCAAAAGATGC  
ATTCAAAATCTTGTTCCAAGAACTCAAGTCACTTTACATCTATTAAGTCTTTGGAGACTTCATAA  
TTTTTCTAGGAAAGGTGTTAGTGGTGTGTTTCACTGTTTTTGGAGGACTCATGGCTTTTAACTAC  
AATCGGGCATTCCAGGTGTGGGCAGTCCCTCTGTTATTGGTAGCTTTTTTTGCCTACTTAGTAGC  
CCATAGTTTTTTATCTGTGTTTGAAACTGTGCTGGATGCACTTTTCCTGTGTTTTGCTGTTGATC  
TGGAACAAATGATGGATCGTCAGAAAAGCCCTACTTTATGGATCAAGAATTTCTGAGTTTCGTA  
AAAAGGAGCAACAAATTAAACAATGCAAGGGGCACAGCAGGACAAGCACTCATTAAGGAATGAGGA  
GGGAACAGAACTCCAGGCCATTGTGAGAT**TAG**ATACCCATTTAGGTATCTGTACCTGGAAAACATT  
TCCTTCTAAGAGCCATTTACAGAATAGAAGATGAGACCACTAGAGAAAAGTTAGTGAATTTTTTT  
TTAAAGACCTAATAAACCTATTCTTCCTCAAAA

## **FIGURE 32**

MSGRDTILGLCILALALSLAMMFTFRFITLLVHIFISLVILGLLFVCGVLWWLYDYDNDLSIE  
LDTERENMKCVLGFAIVSTGITAVLLVLIFVLRKRIKLTVELFQITNKAISSAPFLLFQPLWTF  
ILIFFWVLWVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFILACQQMTIA  
GAVVTCYFNRSKNDPPDHPILSSLSILFFYHQGTVVKGSEFLISVVRIPRIIVMYMQNALKEQQHG  
ALSRYLFRCCYCCFWCLDKYLLHLNQNAYTTTAINGTDFCTSAKDAFKILSKNSSHFTSINCFGD  
FIIFLGKVLVVCFTVFGGLMAFNYNRAFQVWAVPLLLVAFFAYLVAHSFLSVFETVLDALFLCFA  
VDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRNEEGTELQAIVR

**Important features:**

**Signal peptide:**

amino acids 1-20

**Putative transmembrane domains:**

amino acids 35-54, 75-97, 126-146, 185-204, 333-350, 352-371

**N-glycosylation sites.**

amino acids 204-208, 295-299, 313-317

**N-myristoylation sites.**

amino acids 147-153, 178-184, 196-202, 296-275, 342-348

[illegible]

**FIGURE 34**

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPFKFTVPQINCDVKAGKIIDPEFIVKCPAG  
CQDPKYHVYGTDVYASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYSNGVQSLSLPRWR  
ESFIVLESKPKKGVITYPSALTYSSSKSPAAQAGETTKAYQRPPIPGTTAQPVTLMQLLAVTVAVA  
TPTTLPRPSPSAASTTSIPRPQSVGHRSEQEMDLWSTATYTSSQNRPRADPGIQRQDPSGAAFQKP  
VGADVSLGLVPKEELSTQSLEPVSLGDPNCKIDL SFLIDGSTSIGKRRFRIQKQLLADVAQALDI  
GPAGPLMGVVQYGDNPATHFNLKHTNSRDLKTAIEKITQRGGLSNVGRAISFVTKNFFSKANGN  
RSGAPNVVVVMVDGWPTDKVEEASRLARESGINIFFITIEGAAENEKQYVVEPNFANKAVCRTNG  
FYSLHVQSWFGLHKTLLQPLVKRVCDTDR LACSKTCLNSADIGFVIDGSSSVGTGNFRTVLQFVTN  
LTKEFEISD TDTRIGAVQYTYEQRL EFGFDKYSSKPDILNAIKRVGYWSGGTSTGAAINFALEQL  
FKKSKPNKRKLMILITDGRSYDDVRI PAMAAHLKGVITYAIGVAWAAQEELEVIATHPARDHSFF  
VDEFDNLHQYVPRIIQNICTEFNSQPRN

**Important features:****Signal peptide:**

amino acids 1-26

**Transmembrane domain:**

amino acids 181-200

**N-glycosylation sites.**

amino acids 390-394, 520-524

**N-myristoylation sites.**amino acids 23-29, 93-99, 115-121, 262-268, 367-373, 389-395,  
431-437, 466-472, 509-515, 570-576, 571-577, 575-581, 627-633**Amidation site.**

amino acids 304-308

**FIGURE 35**

CCGAGCACAGGAGATTGCCTGCGTTTAGGAGGTGGCTGCGTTGTGGGAAAAGCTATCAAGGAAGAAATTGC  
CAAACCATGTCTTTTTTTCTGTTTTTCAGAGTAGTTTCACAACAGATCTGAGTGTTTTAATTAAGCATGGAAT  
ACAGAAAACAACAAAAAACTTAAGCTTTAATTTTCATCTGGAATTCACAGTTTTCTTAGCTCCCTGGACCC  
GGTTGACCTGTTGGCTCTTCCCGCTGGCTGCTCTATCACGTGGTGCTCTCCGACTACTCACCCCGAGTGTA  
AAGAACCTTCGGCTCGCGTGCTTCTGAGCTGCTGTGGATGGCCTCGGCTCTCTGGACTGTCCTTCCGAGTA  
GGATGTCACCTGAGATCCCTCAAATGGAGCCTCCTGCTGCTGTCACTCCTGAGTTTCTTTGTGATGTGGTAC  
CTCAGCCTTCCCCACTACAATGTGATAGAACGCGTGAAGTGGATGTACTTCTATGAGTATGAGCCGATTTA  
CAGACAAGACTTTCACTTCACACTTCGAGAGCATTCAAAGTCTCTCATCAAAATCCATTTCTGGTCATTC  
TGGTGACCTCCCACCTTCAGATGTGAAAGCCAGGCAGGCCATTAGAGTTACTTGGGGTGAAAAAAGTCT  
TGGTGGGGATATGAGGTTCTTACATTTTTCTTATTAGGCCAAGAGGCTGAAAAGGAAGACAAAATGTTGGC  
ATTGTCCTTAGAGGATGAACACCTTCTTTATGGTGACATAATCCGACAAGATTTTTTAGACACATATAATA  
ACCTGACCTTGAAAACCATTATGGCATTTCAGGTGGGTAACTGAGTTTGGCCCAATGCCAAGTACGTAATG  
AAGACAGACACTGATGTTTTTCATCAATACTGGCAATTTAGTGAAGTATCTTTTAAACCTAAACCACTCAGA  
GAAGTTTTTTCACAGGTTATCCTCTAATTGATAATTATTCCTATAGAGGATTTTACCAAAAAACCCATATTT  
CTTACCAGGAGTATCCTTTCAAGGTGTTCCCTCCATACTGCAGTGGGTGGGTATATAATGTCCAGAGAT  
TTGGTGCCAAGGATCTATGAAATGATGGGTACGTAAAACCCATCAAGTTTGAAGATGTTTATGTCTGGGAT  
CTGTTTGAATTTATTAAAAGTGAACATTCATATTCCAGAAGACACAAATCTTTTCTTCTATATAGAATCC  
ATTTGGATGTCTGTCAACTGAGACGTGTGATTGCAGCCCATGGCTTTTCTTCCAAGGAGATCATCACTTTT  
TGCGAGGTCATGCTAAGGAACACCACATGCCATTATTAACTTCACATTCTACAAAAAGCCTAGAAGGACAG  
GATACCTTGTGGAAAGTGTTAAATAAAGTAGGTACTGTGGAAAATTCATGGGGAGGTCAGTGTGCTGGCTT  
ACACTGAACTGAACTCATGAAAAACCCAGACTGGAGACTGGAGGGTTACACTTGTGATTTATTAGTCAGG  
CCCTTCAAAGATGATATGTGGAGGAATTAAATATAAAGGAATTGGAGGTTTTTGCTAAAGAAATTAATAGG  
ACCAAACAATTTGGACATGTCATTCTGTAGACTAGAATTTCTTAAAGGGTGTTACTGAGTTATAAGCTCA  
CTAGGCTGTAAAAACAAAACAATGTAGAGTTTTTATTTATTGAACAATGTAGTCACTTGAAGGTTTTGTGTA  
TATCTTATGTGGATTACCAATTTAAAAATATATGTAGTTCTGTGTCAAAAACTTCTTCACTGAAGTTATA  
CTGAACAAAATTTTACCTGTTTTTGGTCATTTATAAAGTACTTCAAGATGTTGCAGTATTTACAGTTATT  
ATTATTTAAATTAATTTCACTTTGTGTTTTTAAATGTTTTGACGATTTCAATACAAGATAAAAAAGGATAG  
TGAATCATTTCTTTACATGCAAACATTTTCCAGTTACTTAACTGATCAGTTTATTATTGATACATCACTCCA  
TTAATGTAAAGTCATAGGTCATTATTGCATATCAGTAATCTCTTGGACTTTGTAAATATTTTACTGTGGT  
AATATAGAGAAGAATTAAAGCAAGAAAATCTGAAAA

## **FIGURE 36**

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPYRQDFHF  
TLREHSNCSHQNPFLVILVTSHPSDVKARQAIRVTWGEKKSWWGYEVLTFLLGQEAEKEDKMLA  
LSLEDEHLLYGDIIRQDFLDTYNNLTLKTIMAFRWVTEFCPNAKYVMKTDTDVFINTGNLVKYL  
NLNHSEKFFTGYPLIDNYSYRGFYQKTHISYQYEPFKVFPPYCSGLGYIMSRDLVPRIYEMMGHV  
KPIKFEDVYVGICLNLLKVNIIHIPEDTNLFFLYRIHL DVCQLRRVIAAHGFSSKEIITFWQVMLR  
NTTCHY

**Important features:**

**Type II transmembrane domain:**

amino acids 20-39

**N-glycosylation sites.**

amino acids 72-76, 154-158, 198-202, 212-216, 326-330

**Glycosaminoglycan attachment site.**

amino acids 239-243

**Ly-6 / u-PAR domain proteins.**

amino acids 23-37

**N-myristoylation site.**

amino acids 271-277



### FIGURE 37

[illegible]

## **FIGURE 38**

MELGCWTQLGLTFLQLLLISSLPREYTVINEACPGAENIMCRECCEYDQIECVCPGKREVVGYT  
 IPCCRNEENECDSCLIHPGCTIFENCKSCRNGSWGGLDDEFYVKGIFYCAECRAGWYGGDCMRCGQ  
 VLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRDGDNRDGGQII  
 KRVCGNERPAPIQSIGSSLHVLFSHDGSKNFDGFHAIYEEITACSSSPCFHDGTCVLDKAGSYKC  
 ACLAGYTGQRCENLLEERNCSDPGGPVNGYQKITGGPGLINGRHAKIGTVVSFFCNNSYVLSGNE  
 KRTCQONGEWSGKQPICIKACREPKISDLVRRRVLPMPVQSRETPLHQLYSAAFSKQKLQSAPTK  
 KPALPFGDLPMGYQHLHTQLQYECISPFYRRLGSSRRTCLRTGKWSGRAPSCIPICGKIENITAP  
 KTQGLRWPWQAAIYRRTSGVHDGSLHKGAWFLVCSGALVNERTVVVAAHCVTDLGKVTMIKTADL  
 KVLGKFYRDDDRDEKTIQSLQISAILHPNYDPILLDADIAILKLLDKARISTRVQPICLAASR  
 DLSTSFQESHITVAGWNVLADVRSPGFKNDTLRSGVVSVVDSLLCEEQHEDHGIPVSVTDNMFCA  
 SWEPTAPSDICTAETGGIAAVSFPGRASPEPRWHLMGLVSWSYDKTCSHRLSTAFTKVLPFKDWI  
 ERNMK

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-23

#### **EGF-like domain cysteine pattern signature.**

amino acids 260-272

#### **N-glycosylation sites.**

amino acids 96-100, 279-283, 316-320, 451-455, 614-618

#### **N-myristoylation sites.**

amino acids 35-41, 97-103, 256-262, 284-290, 298-304, 308-314,  
 474-480, 491-497, 638-644, 666-672

#### **Amidation site.**

amino acids 56-60

#### **Serine proteases, trypsin family.**

amino acids 489-506

#### **CUB domain proteins profile.**

amino acids 150-167

**FIGURE 39**

GGTTCCTACATCCTCTCATCTGAGAATCAGAGAGCATAATCTTCTTACGGGCCCCGTGATTTATTAACGTGGCTTAATC  
TGAAGGTTCTCAGTCAAATTCTTTGTGATCTACTGATTGTGGGGGCATGGCAAGGTTTGCTTAAAGGAGCTTGGCTGG  
TTTGGGCCCTTGTAGCTGACAGAAGGTGGCCAGGGAGAATGCAGCACACTGCTCGGAGAATGAAGGCGCTTCTGTTGC  
TGGTCTTGGCCTTGGCTCAGTCCTGCTAACTACATTGACAATGTGGGCAACCTGCACTTCCTGTATTGAGAACTCTGTA  
AAGGTGCCTCCCACTACGGCCTGACCAAAGATAGGAAGAGGCGCTCACAAGATGGCTGTCCAGACGGCTGTGCGAGCC  
TCACAGCCACGGCTCCCTCCCCAGAGGTTTCTGCAGCTGCCACCATCTCCTTAATGACAGACGAGCCTGGCCTAGACA  
ACCCTGCCTACGTGTCTCTCGGCAGAGGACGGGCAGCCAGCAATCAGCCCAGTGGACTCTGGCCGGAGCAACCGAACTA  
GGGCACGGGCCCTTTGAGAGATCCACTATTAGAAGCAGATCATTTAAAAAATAAATCGAGCTTTGAGTGTCTCTCGAA  
GGACAAAGAGCGGGAGTGCAGTTGCCAACCATGCCGACCAGGGCAGGGAAAATTCTGAAAACACCACTGCCCCCTGAAG  
TCTTTCCAAGGTTGTACCACCTGATTCCAGATGGTGAAATTACCAGCATCAAGATCAATCGAGTAGATCCCAGTGAAA  
GCCTCTCTATTAGGCTGGTGGGAGGTAGCGAAACCCCACTGGTCCATATCATTATCCAACACATTTATCGTGATGGGG  
TGATCGCCAGAGACGGCCGGCTACTGCCAGGAGACATCATTCTAAAGGTCAACGGGATGGACATCAGCAATGTCCCTC  
ACAACTACGCTGTGCGTCTCTGCGGCAGCCCTGCCAGGTGCTGTGGCTGACTGTGATGCGTGAACAGAAGTTCCGCA  
GCAGGAACAATGGACAGGCCCCGGATGCCTACAGACCCCGAGATGACAGCTTTCATGTGATTCTCAACAAAAGTAGCC  
CCGAGGAGCAGCTTGAATAAAACTGGTGCGCAAGGTGGATGAGCCTGGGGTTTTTCATCTTCAATGTGCTGGATGGCG  
GTGTGGCATATCGACATGGTCAGCTTGAGGAGAATGACCGTGTGTTAGCCATCAATGGACATGATCTTCGATATGGCA  
GCCCAGAAAGTGCGGCTCATCTGATTGAGGCCAGTGAAAGACGTGTTACCTCGTCGTGTCCCGCCAGGTTCCGGCAGC  
GGAGCCCTGACATCTTTTCAGGAAGCCGGCTGGAACAGCAATGGCAGCTGGTCCCCAGGGCCAGGGGAGAGGAGCAACA  
CTCCCAAGCCCCTCCATCTACAATTACTTGTGTCATGAGAAGGTGGTAAATATCCAAAAGACCCCGGTGAATCTCTCG  
GCATGACCGTCGCAGGGGGAGCATCACATAGAGAATGGGATTTGCCTATCTATGTCATCAGTGTTGAGCCCGGAGGAG  
TCATAAGCAGAGATGGAAGAATAAAACAGGTGACATTTTGTGTAATGTGGATGGGGTGAAGTACAGAGGTCAGCC  
GGAGTGAGGCAGTGGCATTATTGAAAAGAACATCATCCTCGATAGTACTCAAAGCTTTGGAAGTCAAAGAGTATGAGC  
CCCAGGAAGACTGCAGCAGCCAGCAGCCCTGGACTCCAACCACAACATGGCCCCACCCAGTGACTGGTCCCCATCCT  
GGGTCATGTGGCTGGAATTACCACGGTGCTTGTATAACTGTAAAGATATTGTATTACGAAGAAACACAGCTGGAAGTC  
TGGGCTTCTGCATTGTAGGAGGTTATGAAGAATACAATGGAAACAAACCTTTTTTTCATCAAATCCATTGTTGAAGGAA  
CACCAGCATACAATGATGGAAGAATTAGATGTGGTGATATTCTTCTTGCTGTCAATGGTAGAAGTACATCAGGAATGA  
TACATGCTTGCTTGGCAAGACTGCTGAAAGAACTTAAAGGAAGAATTACTCTAACTATTGTTTCTTGGCCTGGCACTT  
TTTTATAGAAATCAATGATGGGTGAGAGGAAAACAGAAAAATCACAAATAGGCTAAGAAGTTGAAACACTATATTTATC  
TTGTGAGTTTTTATATTTAAAGAAAGAATACATTGTAAAAATGTCAGGAAAAGTATGATCATCTAATGAAAGCCAGTT  
ACACCTCAGAAAATATGATTCCAAAAAATTAATACTACTAGTTTTTTTTTTCAGTGTGGAGGATTTCTCATTACTCTAC  
AACATTGTTTATATTTTTTCTATTCAATAAAAAGCCCTAAAACAATAAATGATTGATTTGTATACCCCACTGAATT  
CAAGCTGATTTAAATTTAAATTTGGTATATGCTGAAGTCTGCCAAGGGTACATTATGGCCATTTTTTAATTTACAGCT  
AAAATATTTTTTAAATGCATTGCTGAGAAACGTTGCTTTCATCAAACAAGAATAAATATTTTTTCAGAAGTTAAA

## **FIGURE 40**

MKALLLLVLPWLSPANYIDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGCPCDGCASLTATAPS  
PEVSAAATISLMTDEPGLDNPAYVSSAEDGQPAISPVDSGRSNRTRARPFERSTIRSRSEFKKINR  
ALSVLRRTKSGSAVANHADQGRENSENTTAPEVFPRLYHLIPDGEITSIKINRVDPSESLSIRLV  
GGSETPLVHIIIIQHIYRDGVIARDGRLLPGDIILKVNGMDISNVPHNYAVRLLRQPCQVLWLTVM  
REQKFRSRNNGQAPDAYRPRDDSFHVILNKSSPEEQLGIKLVKRVDEPGVFIFNVLDGGVAYRHG  
QLEENDRVLAINGHDLRYGSPESA AHLIQA SERRVHLVVS RQVRQ RSPDIFQEAGWNSNGSWSPG  
PGRSNTPKPLHPTITCHEKV VNIQKDPGESLGMTVAGGASHREWDLPIYVISVEPGGVISRDGR  
IKTGDILLNVDGVELTEVSRSEAVALLKRTSSSIVLKALEVKEYEPQEDCSSPAALDSNHNMAPP  
SDWSPSWVMWLELPRCLYNCKDIVLRRNTAGSLGFCIVGGYEEYNGNKPFFIKSIVEGTPAYNDG  
RIRCGDILLAVNGRSTSGMIHACLARLLKELKGRITLTIVSWPGTFL

**Important features:**

**Signal peptide:**

amino acids 1-15

**N-glycosylation sites.**

amino acids 108-112, 157-161, 289-293, 384-388

**Tyrosine kinase phosphorylation sites.**

amino acids 433-441, 492-500

**N-myristoylation sites.**

amino acids 51-57, 141-147, 233-239, 344-350, 423-429, 447-453,  
467-473, 603-609

**FIGURE 41**

ACCAGGCATTGTATCTTCAGTTGTCATCAAGTTCGCAATCAGATTGGAAAAGCTCAACTTGAAGC  
TTTCTTGCCTGCAGTGAAGCAGAGAGATAGATATTATTACGTAATAAAAAAC**ATG**GGCTTCAAC  
CTGACTTTCCACCTTTCCCTACAAATTCCGATTACTGTTGCTGTTGACTTTGTGCCTGACAGTGGT  
TGGGTGGGCCACCAGTAACTACTTCGTGGGTGCCATTCAAGAGATTCCCTAAAGCAAAGGAGTTCA  
TGGCTAATTTCCATAAGACCCTCATTTTGGGGAAGGGAAAACTCTGACTAATGAAGCATCCACG  
AAGAAGGTAGAACTTGACAACCTGTCCTTCTGTGTCTCCTTACCTCAGAGGCCAGAGCAAGCTCAT  
TTTCAAACCAGATCTCACTTTTGGAAAGAGGTACAGGCAGAAAATCCCAAAGTGTCCAGAGGCCGGT  
ATCGCCCTCAGGAATGTAAAGCTTTACAGAGGGTCGCCATCCTCGTTCCCCACCGGAACAGAGAG  
AAACACCTGATGTACCTGCTGGAACATCTGCATCCCTTCCTGCAGAGGCAGCAGCTGGATTATGG  
CATCTACGTCATCCACCAGGCTGAAGGTAAAAAGTTTAATCGAGCCAACTCTTGAATGTGGGCT  
ATCTAGAAGCCCTCAAGGAAGAAAATTGGGACTGCTTTATATTCCACGATGTGGACCTGGTACCC  
GAGAATGACTTTAACCTTTACAAGTGTGAGGAGCATCCCAAGCATCTGGTGGTTGGCAGGAACAG  
CACTGGGTACAGGTTACGTTACAGTGGATATTTTGGGGGTGTTACTGCCCTAAGCAGAGAGCAGT  
TTTTCAAGGTGAATGGATTCTCTAACAACCTACTGGGGATGGGGAGGCGAAGACGATGACCTCAGA  
CTCAGGGTTGAGCTCCAAAGAATGAAAATTTCCCGGCCCTGCCTGAAGTGGGTAAATATACAAT  
GGTCTTCCACACTAGAGACAAAGGCAATGAGGTGAACGCAGAACGGATGAAGCTCTTACACCAAG  
TGTCACGAGTCTGGAGAACAGATGGGTTGAGTAGTTGTTCTTATAAATTAGTATCTGTGGAACAC  
AATCCTTTATATATCAACATCACAGTGGATTTCTGGTTTGGTGCA**TGA**CCCTGGATCTTTTGGTG  
ATGTTTGAAGAAGCTGATTCTTTGTTTGCAATAATTTTGGCCTAGAGACTTCAAATAGTAGCACA  
CATTAAAGAACCTGTTACAGCTCATTGTTGAGCTGAATTTTTCCTTTTGTATTTTCTTAGCAGAG  
CTCCTGGTGATGTAGAGTATAAAACAGTTGTAACAAGACAGCTTCTTAGTCATTTTGATCATGA  
GGGTAAATATTGTAATATGGATACTTGAAGGACTTTATATAAAAGGATGACTCAAAGGATAAAA  
TGAACGCTATTTGAGGACTCTGGTTGAAGGAGATTTATTTAAATTTGAAGTAATATATTATGGGA  
TAAAAGGCCACAGGAAATAAGACTGCTGAATGTCTGAGAGAACCAGAGTTGTTCTCGTCCAAGGT  
AGAAAGGTACGAAGATACAATACTGTTATTCAATTTATCCTGTACAATCATCTGTGAAGTGGTGGT  
GTCAGGTGAGAAGGCGTCCACAAAAGAGGGGAGAAAAGGCGACGAATCAGGACACAGTGAACCTTG  
GGAATGAAGAGGTAGCAGGAGGGTGGAGTGTGGCTGCAAAGGCAGCAGTAGCTGAGCTGGTTGC  
AGGTGCTGATAGCCTTCAGGGGAGGACCTGCCCAGGTATGCCTTCCAGTGATGCCCACCAGAGAA  
TACATTCTCTATTAGTTTTTAAAGAGTTTTTGTAAAATGATTTTGTACAAGTAGGATATGAATTA  
GCAGTTTACAAGTTTACATATTAATAATAATAATATGTCTATCAAATACCTCTGTAGTAAAT  
GTGAAAAGCAAAA

## **FIGURE 42**

MGFNLT FHL SYKFRLLLLLTLCLTVVGWATSNYFVGAIQEIPKAKEFMANFHKTILGKGKTLTN  
EASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRVAILVPH  
RNREKHLMYLLEHLHPFLQRQQLDYGIYVIHQAEKKFNRAKLLNVGYLEALKEENWDCFI FHDV  
DLVPENDFNLYKCEEHPKHLVVGRNSTGYRLRYSGYFGGVTALSREQFFKVNGFSNNYWG WGGED  
DDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSRVWR TDGLSSCSYKLV  
SVEHNPLYINITVDFWFGA

**Important features:**

**Signal peptide:**

amino acids 1-27

**N-glycosylation sites.**

amino acids 4-8, 220-224, 335-339

**Xylose isomerase proteins.**

amino acids 191-202

**FIGURE 43**

GCTCAAGACCCAGCAGTGGGACAGCCAGACAGACGGCACG**ATG**GCACTGAGCTCCCAGATCTGGG  
CCGCTTGCCTCCTGCTCCTCCTCCTCCTCGCCAGCCTGACCAGTGGCTCTGTTTTCCCAACAACAG  
ACGGGACAACCTTGCAGAGCTGCAACCCCAGGACAGAGCTGGAGCCAGGGCCAGCTGGATGCCCAT  
GTTCCAGAGGCGAAGGAGGCGAGACACCCACTTCCCCATCTGCATTTTCTGCTGCGGCTGCTGTC  
ATCGATCAAAGTGTGGGATGTGCTGCAAGACG**TAG**AACCTACCTGCCCTGCCCCGTCCCCTCCC  
TTCCTTATTTATTCCTGCTGCCCCAGAACATAGGTCTTGGAATAAAATGGCTGGTTCTTTTGTTT  
TCCAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 44**

MALSSQIWAACLLLLLLLLASLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRRDTHFPI  
CIFCCGCCCHRSKCGMCCKT

**Important features:**

**Signal peptide:**

amino acids 1-24

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 58-59

**N-myristoylation site.**

amino acids 44-50

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 1-12



**FIGURE 45**

GTGGCTTCATTTTCAGTGGCTGACTTCCAGAGAGCAAT**ATG**GCTGGTTCCCCAACATGCCTCACCC  
TCATCTATATCCTTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTGGTCGGT  
TCCGTTGGTGGGGCCGTGACTTTCCCCCTGAAGTCCAAAGTAAAGCAAGTTGACTCTATTGTCTG  
GACCTTCAACACAACCCCTCTTGTCACCATAACAGCCAGAAGGGGGCACTATCATAGTGACCCAAA  
ATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAGCTCAGCAAACCTGAAG  
AAGAATGACTCAGGGATCTACTATGTGGGGATATACAGCTCATCACTCCAGCAGCCCTCCACCCA  
GGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAGTCACCATGGGTCTGCAGAGCA  
ATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATGGAACATGGGGAAGAGGATGTGATT  
TATACCTGGAAGGCCCTGGGGCAAGCAGCCAATGAGTCCCATAATGGGTCCATCCTCCCCATCTC  
CTGGAGATGGGGAGAAAGTGATATGACCTTCATCTGCGTTGCCAGGAACCCTGTCAGCAGAACT  
TCTCAAGCCCCATCCTTGCCAGGAAGCTCTGTGAAGGTGCTGCTGATGACCCAGATTCCCTCCATG  
GTCCTCCTGTGTCTCCTGTTGGTGCCCCCTCCTGCTCAGTCTCTTTGTACTGGGGCTATTTCTTTG  
GTTTCTGAAGAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAAGAGAGTGGACATTTGTCGGG  
AAACTCCTAACATATGCCCCCATCTTGAGAGAAACACAGAGTACGACACAATCCCTCACACTAAT  
AGAACAATCCTAAAGGAAGATCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAAAGAT  
GGAAAATCCCCACTCACTGCTCACGATGCCAGACACACCAAGGCTATTTGCCTATGAGAATGTTA  
TC**TAG**ACAGCAGTGCACTCCCCTAAGTCTCTGCTCA

## **FIGURE 46**

MAGSPTCLTLIYILWQLTGSAASGPVKELVGSVGGAVTFPLKSKVKQVDSIVWTFNTTPLVTIQP  
EGGTIIIVTQNRNRERVDFFPDGGYSLKLSKLKKNDSGIYYVGIYSSSLQQPSTQEYVLHVYEHLSK  
PKVTMGLQSNKNGTCVTNLTCCMEHGEEDVIYTWKALGQAANESHNGSILPISWRWGESDMTFIC  
VARNPVSRNFSSPILARKLCEGAADDPDSSMVLLCLLLVPLLLSLFVLGLFLWFLKRERQEEYIE  
EKKRVDICRETPNICPHSGENTHEYDTIPHTNRTILKEDPANTVYSTVEIPKKMENPHSLLTMPDT  
PRLFAYENVI

**Important features:**

**Signal peptide:**

amino acids 1-22

**Transmembrane domain:**

amino acids 224-250

**Leucine zipper pattern.**

amino acids 229-251

**N-glycosylation sites.**

amino acids 98-102, 142-146, 148-152, 172-176, 176-180, 204-208,  
291-295

**FIGURE 47**

GGCTCGAGCGTTTCTGAGCCAGGGGTGACC**ATG**ACCTGCTGCGAAGGATGGACATCCTGCAATGG  
ATTGAGCCTGCTGGTTCTACTGCTGTTAGGAGTAGTTCTCAATGCGATACCTCTAATTGTCAGCT  
TAGTTGAGGAAGACCAATTTTCTCAAACCCCATCTCTTGCTTTGAGTGGTGGTTCCCAGGAATT  
ATAGGAGCAGGTCTGATGGCCATTCCAGCAACAACAATGTCCTTGACAGCAAGAAAAAGAGCGTG  
CTGCAACAACAGAACTGGAATGTTTCTTTCATCATTTTTTCAGTGTGATCACAGTCATTGGTGCTC  
TGTATTGCATGCTGATATCCATCCAGGCTCTCTTAAAAGGTCCTCTCATGTGTAATTCTCCAAGC  
AACAGTAATGCCAATTGTGAATTTTCATTGAAAAACATCAGTGACATTCATCCAGAATCCTTCAA  
CTTGCAAGTGGTTTTTCAATGACTCTTGTGCACCTCCTACTGGTTTTCAATAAACCCACCAGTAACG  
ACACCATGGCGAGTGGCTGGAGAGCATCTAGTTTCCACTTCGATTCTGAAGAAAACAAACATAGG  
CTTATCCACTTCTCAGTATTTTTTAGGTCTATTGCTTGTGGAATTCTGGAGGTCCTGTTTGGGCT  
CAGTCAGATAGTCATCGGTTTCCTTGGCTGTCTGTGTGGAGTCTCTAAGCGAAGAAGTCAAATTG  
TG**TAG**TTTAATGGGAATAAAATGTAAGTATCAGTAGTTTGAAAAAAAAA

## **FIGURE 48**

MTCCEGWTSCNGFSLLVLLLLGVVLNAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMAIPA  
TTMSLTARKRACCNNRTGMFLSSFFSVITVIGALYCMLISIQALLKGPLMCNSPSNSNANCEFSL  
KNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHFDSEENKHRLIHFSVFLGL  
LLVGILEVLFGLSQIVIGFLGCLCGVSKRRSQIV

**Important features:**

**Transmembrane domains:**

amino acids 10-31 (type II), 50-72, 87-110, 191-213

**N-glycosylation sites.**

amino acids 80-84, 132-136, 148-152, 163-167

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 223-227

**N-myristoylation sites.**

amino acids 22-28, 54-60, 83-89, 97-103, 216-222

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 207-218

**TNFR/NGFR family cysteine-rich region protein.**

amino acids 4-12

**FIGURE 49**

ATCCGTTCTCTGCGCTGCCAGCTCAGGTGAGCCCTCGCCAAGGTGACCTCGCAGGACACTGGTGA  
AGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTGGAGCAG  
ATCCGTGGGCTGCAGACCCCCGCCCCAGTGCCTCTCCCCCTGCAGCCCTGCCCCCTCGAACTGTGA  
**CATG**GAGAGAGTGACCCTGGCCCTTCTCCTACTGGCAGGCCTGACTGCCTTGGAAGCCAATGACC  
CATTTGCCAATAAAGACGATCCCTTCTACTATGACTGGAAAAACCTGCAGCTGAGCGGACTGATC  
TGCGGAGGGCTCCTGGCCATTGCTGGGATCGCGGCAGTTCTGAGTGGCAAATGCAAATACAAGAG  
CAGCCAGAAGCAGCACAGTCCTGTACCTGAGAAGGCCATCCCACTCATCACTCCAGGCTCTGCCA  
CTACTTGCT**TGA**GCACAGGACTGGCCTCCAGGGATGGCCTGAAGCCTAACACTGGCCCCCAGCACC  
TCCTCCCCCTGGGAGGCCTTATCCTCAAGGAAGGACTTCTCTCCAAGGGCAGGCTGTTAGGCCCT  
TTCTGATCAGGAGGCTTCTTTATGAATTAACTCGCCCCACCACCCCTCA

## **FIGURE 50**

MERVTLALLLLAGLTALEANDPFANKDDPFYYDWKNLQLSGLICGGLLAIAGIAAVLSGKCKYKS  
SQKQHSPVPEKAIPPLITPGSATTC

**Important features:**

**Signal peptide:**

amino acids 1-16

**Transmembrane domain:**

amino acids 36-59

**N-myristoylation sites.**

amino acids 41-47, 45-51, 84-90

**Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7.**

amino acids 54-67

[illegible]

## **FIGURE 52**

MKFQGPLACLLLLALCLGSGEAGPLQSGEESTGTNIGEALGHGLGDALSEGVGKAIGKEAGGAAGS  
KVSEALGQGTREAVGTGVRQVPFGGAADALGNRVGEAAHALGNTGHEIGRQAEDVIRHGADAVRG  
SWQGVPGHSGAWETSGGHGIFGSQGGGLGGQGNPGGLGTPWVHGYPGNSAGSFGMNPQGAPWGQ  
GGNGGPPNFGTNTQGAQAQPGYGSVRASNQNEGCTNPPPSGSGGGSSNSGGGSGSQSGSSGSGSN  
GDNNNGSSSGSSSGSSSGSSSGSSSGSSSGSSSGNSGGSRGDSGSESSWGSSTGSSSGNHGGSG  
GGNGHKPGCEKPGNEARGSGESGIQGFRGQGVSSNMREISKEGNRLLGGSGDNRYRGQSSWGS  
GDAVGGVNTVNSETSPGMFNFDTFWKNFKSKLGFINWDAINKDQRSSRIP

**Signal peptide:**

amino acids 1-21

**N-glycosylation site.**

amino acids 265-269

**Glycosaminoglycan attachment site.**

amino acids 235-239, 237-241, 244-248, 255-259, 324-328, 388-392

**Casein kinase II phosphorylation site.**

amino acids 26-30, 109-113, 259-263, 300-304, 304-308

**N-myristoylation site.**

amino acids 17-23, 32-38, 42-48, 50-56, 60-66, 61-67, 64-70, 74-80,  
90-96, 96-102, 130-136, 140-146, 149-155, 152-158, 155-161, 159-165,  
163-169, 178-184, 190-196, 194-200, 199-205, 218-224, 236-242,  
238-244, 239-245, 240-246, 245-251, 246-252, 249-252, 253-259,  
256-262, 266-272, 270-276, 271-277, 275-281, 279-285, 283-289,  
284-290, 287-293, 288-294, 291-297, 292-298, 295-301, 298-304,  
305-311, 311-317, 315-321, 319-325, 322-328, 323-329, 325-331,  
343-349, 354-360, 356-362, 374-380, 381-387, 383-389, 387-393,  
389-395, 395-401

**Cell attachment sequence.**

amino acids 301-304



**FIGURE 53**

GGAGAAGAGGTTGTGTGGGACAAGCTGCTCCCGACAGAAGG**ATG**TCGCTGCTGAGCCTGCCCTGG  
CTGGGCCTCAGACCGGTGGCAATGTCCCCATGGCTACTCCTGCTGCTGGTTGTGGGCTCCTGGCT  
ACTCGCCCGCATCCTGGCTTGGACCTATGCCTTCTATAACAACCTGCCGCCGGCTCCAGTGTTTCC  
CACAGCCCCCAAACGGAACCTGGTTTTGGGGTCACCTGGGCCTGATCACTCCTACAGAGGAGGGC  
TTGAAGGACTCGACCCAGATGTGCGCCACCTATTCCCAGGGCTTTACGGTATGGCTGGGTCCCAT  
CATCCCCTTCATCGTTTTATGCCACCCTGACACCATCCGGTCTATCACCAATGCCTCAGCTGCCA  
TTGCACCCAAGGATAATCTCTTCATCAGGTTCTGAAGCCCTGGCTGGGAGAAGGGATACTGCTG  
AGTGGCGGTGACAAGTGGAGCCGCCACCGTCGGATGCTGACGCCCGCCTTCCATTTCAACATCCT  
GAAGTCCTATATAACGATCTTCAACAAGAGTGCAAACATCATGCTTGACAAGTGGCAGCACCTGG  
CCTCAGAGGGCAGCAGTCGTCTGGACATGTTTGAGCACATCAGCCTCATGACCTTGGACAGTCTA  
CAGAAATGCATCTTCAGCTTTGACAGCCATTGTCAGGAGAGGCCCAGTGAATATATTGCCACCAT  
CTTGAGCTCAGTGCCCTTGTAGAGAAAAGAAGCCAGCATATCCTCCAGCACATGGACTTTTCTGT  
ATTACCTCTCCCATGACGGGCGGCGCTTCCACAGGGCCTGCCGCCTGGTGCATGACTTCACAGAC  
GCTGTCATCCGGGAGCGGCGTCGCACCCTCCCCACTCAGGGTATTGATGATTTTTTTCAAAGACAA  
AGCCAAGTCCAAGACTTTGGATTTCAATTGATGTGCTTCTGCTGAGCAAGGATGAAGATGGGAAGG  
CATTGTCAGATGAGGATATAAGAGCAGAGGCTGACACCTTCATGTTTGGAGGCCATGACACCACG  
GCCAGTGGCCTCTCCTGGGTCTGTACAACCTTGCGAGGCACCCAGAATACCAGGAGCGCTGCCG  
ACAGGAGGTGCAAGAGCTTCTGAAGGACCGCGATCCTAAAGAGATTGAATGGGACGACCTGGCCC  
AGCTGCCCTTCCTGACCATGTGCGTGAAGGAGAGCCTGAGGTTACATCCCCCAGCTCCCTTCATC  
TCCCGATGCTGCACCCAGGACATTGTTCTCCAGATGGCCGAGTCATCCCCAAAGGCATTACCTG  
CCTCATCGATATTATAGGGGTCCATCACAACCCAACTGTGTGGCCGGATCCTGAGGTCTACGACC  
CCTTCCGCTTTGACCCAGAGAACAGCAAGGGGAGGTCACCTCTGGCTTTTATTCCTTTCTCCGCA  
GGGCCCAGGAACCTGCATCGGGCAGGCGTTCGCCATGGCGGAGATGAAAGTGGTCCTGGCGTTGAT  
GCTGCTGCACTTCCGGTTCCTGCCAGACCACACTGAGCCCCGCAGGAAGCTGGAATTGATCATGC  
GCGCCGAGGGCGGGCTTTGGCTGCGGGTGGAGCCCCTGAATGTAGGCTTGCA**TGA**CTTTCTGAC  
CCATCCACCTGTTTTTTTTGCAGATTGTCATGAATAAAACGGTGCTGTCAA

## **FIGURE 54**

MSLLSLPWLGLRPVAMSPWLLLLLVVGSWLLARILAWTYAFYNNCRRLQCFPPKRNWFWGHLG  
LITPTEEGLKDSTQMSATYSQGFTVWLGPPIIPFIVLCHPDTIRSITNASAAIAPKDNLFIRFLKP  
WLGEIGILLSGGDKWSRHRMLTPAFHFNILKSYITIFNKSANIMLDKWQHLASEGSSRLDMFEHI  
SLMTLDSLQKCIFSFDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYYLSHDGRRFHRAC  
RLVHDFTDVIRERRRTLPTQGIDFFKDKAKSKTLDFIDVLLLSKDEDGKALSDEDIRAEADTF  
MFGGHDTTASGLSWVLYNLARHPEYQERCQEVQELLKDRDPKEIEWDDLAQLPFLTMCVKESLR  
LHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVWPDPEVYDPFRFDPENSKGRSP  
LAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHTEPRRKLELIMRAEGGLWLRVEPLN  
VGLQ

**Important features:**

**Transmembrane domains:**

amino acids 13-32 (type II), 77-102

**Cytochrome P450 cysteine heme-iron ligand signature.**

amino acids 461-471

**N-glycosylation sites.**

amino acids 112-116, 168-172

## FIGURE 55

[illegible]

## **FIGURE 56**

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCSAFWWHNKGLALIFCILQSLALTWYSLSFIPFAR  
DAVKKCFVCLA

**Important features:**

**Signal peptide:**

amino acids 1-33

**Type II fibronectin collagen-binding domain protein.**

amino acids 30-72

[illegible]

## **FIGURE 58**

MLCLCLYVPVIGEAQTEFQYFESKGLPAELKSIFKLSVFIPSQEFSTYRQWKQKIVQAGDKDLG  
QLDFEEFVHYLQDHEKKRLRVFKILDKKNDGRIDAQEIMQSLRDLGVKISEQQAEKILKSMDKNG  
TMTIDWNEWRDYHLLHPVENIPEIILYWKHSTIFDVGENLTVPDEFTVEERQTGMWWRHLVAGGG  
AGAVSRTCTAPLDRCLKVLMQVHASRSNNMGIVGGFTQMIREGGARSLWRGNGINVLKIAPESAIAK  
FMAYEQIKRLVGSDQETLRIHERLVAGSLAGAIQSSIIYPMEVLKTRMALRKTGQYSGMLDCARR  
ILAREGVAAFYKGYVPNMLGIIIPYAGIDLAVYETLKNWLQHYAVNSADPGVFVLLACGTMSSTC  
QQLASYPLALVRTRMQAQASIEGAPEVTMSSLFKHILRTEGAFGLYRGLAPNFMKVIPAVSISYV  
VYENLKITLGVQSR

**Important features:**

**Signal peptide:**

amino acids 1-16

**Putative transmembrane domains:**

amino acids 284-304, 339-360, 376-394

**Mitochondrial energy transfer proteins signature.**

amino acids 206-215, 300-309

**N-glycosylation sites.**

amino acids 129-133, 169-173

**Elongation Factor-hand calcium-binding protein.**

amino acids 54-73, 85-104, 121-140

**FIGURE 59**

GGAAGGCAGCGGCAGCTCCACTCAGCCAGTACCCAGATACGCTGGGAACCTTCCCCAGCC**ATGGC**  
TTCCCTGGGGCAGATCCTCTTCTGGAGCATAATTAGCATCATCATTATTCTGGCTGGAGCAATTG  
CACTCATCATTGGCTTTGGTATTTTCAGGGAGACACTCCATCACAGTCACTACTGTGCCTCAGCT  
GGGAACATTGGGGAGGATGGAATCCTGAGCTGCACTTTTGAACCTGACATCAAACCTTCTGATAT  
CGTGATACAATGGCTGAAGGAAGGTGTTTTAGGCTTGGTCCATGAGTTCAAAGAAGGCAAAGATG  
AGCTGTCTGGAGCAGGATGAAATGTTTCAGAGGCCGGACAGCAGTGTTTGCTGATCAAGTGATAGTT  
GGCAATGCCTCTTTGCGGCTGAAAAACGTGCAACTCACAGATGCTGGCACCTACAAATGTTATAT  
CATCACTTCTAAAGGCAAGGGGAATGCTAACCTTGAGTATAAACTGGAGCCTTCAGCATGCCGG  
AAGTGAATGTGGACTATAATGCCAGCTCAGAGACCTTGCGGTGTGAGGCTCCCCGATGGTTCCCC  
CAGCCCACAGTGGTCTGGGCATCCCAAGTTGACCAGGGAGCCAACCTTCTCGGAAGTCTCCAATAC  
CAGCTTTGAGCTGAACTCTGAGAATGTGACCATGAAGGTTGTGTCTGTGCTCTACAATGTTACGA  
TCAACAACACATACTCCTGTATGATTGAAAATGACATTGCCAAAGCAACAGGGGATATCAAAGTG  
ACAGAATCGGAGATCAAAGGCGGAGTCACCTACAGCTGCTAAACTCAAAGGCTTCTCTGTGTGT  
CTCTTCTTTCTTTGCCATCAGCTGGGCACTTCTGCCTCTCAGCCCTTACCTGATGCTAAAA**TAAT**  
GTGCCTTGGCCACAAAAAAGCATGCAAAGTCATTGTTACAACAGGGATCTACAGAACTATTTTAC  
CACCAGATATGACCTAGTTTTTATATTTCTGGGAGGAAATGAATTCATATCTAGAAGTCTGGAGTG  
AGCAAACAAGAGCAAGAAACAAAAAGAAGCCAAAAGCAGAAGGCTCCAATATGAACAAGATAAAT  
CTATCTTCAAAGACATATTAGAAGTTGGGAAAATAATTCATGTGAACTAGACAAGTGTGTTAAGA  
GTGATAAGTAAAATGCACGTGGAGACAAGTGCATCCCCAGATCTCAGGGACCTCCCCCTGCCTGT  
CACCTGGGGAGTGAGAGGACAGGATAGTGCATGTTCTTTGTCTCTGAATTTTTTAGTTATATGTGC  
TGTAATGTTGCTCTGAGGAAGCCCCTGGAAAGTCTATCCCAACATATCCACATCTTATATTCCAC  
AAATTAAGCTGTAGTATGTACCCTAAGACGCTGCTAATTGACTGCCACTTCGCAACTCAGGGGCG  
GCTGCATTTTATAGTAATGGGTCAAATGATTCACCTTTTTATGATGCTTCCAAAGGTGCCTTGGCTTC  
TCTTCCCAACTGACAAATGCCAAAGTTGAGAAAAATGATCATAATTTTAGCATAAACAGAGCAGT  
CGGGGACACCGATTTTATAAATAAACTGAGCACCTTCTTTTTTAAACAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 60**

MASLGQILFWSIIISIIIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDIKLS  
DIVIQWLKEGVLGLVHEFKEGKDELSEQDEMFRGRTAVFADQVIVGNASLRLKNVQLTDAGTYKC  
YIITSKGKGNANLEYKTGAFSMPEVNVDYNASSETLRCEAPRWFPQPTVVWASQVDQGANSFSEVS  
NTSFELNSENVTMKVSVLYNVTINNTYSCMIENDIAKATGDIKVTSEIKRRSHLQLLNSKASL  
CVSSFFAISWALLPLSPYLMLK

**Important features:**

**Signal peptide:**

amino acids 1-28

**Transmembrane domain:**

amino acids 258-281

**N-glycosylation sites.**

amino acids 112-116, 160-164, 190-194, 196-200, 205-209, 216-220,  
220-224

**N-myristoylation sites.**

amino acids 52-58, 126-132, 188-194



**FIGURE 61**

TGACGTCAGAATCACCATGGCCAGCTATCCTTACCGGCAGGGCTGCCCAGGAGCTGCAGGACAAG  
CACCAGGAGCCCCCTCCGGGTAGCTACTACCCTGGACCCCCCAATAGTGGAGGGCAGTATGGTAGT  
GGGCTACCCCCTGGTGGTGGTTATGGGGGTCTGCCCCTGGAGGGCCTTATGGACCACCAGCTGG  
TGGAGGGCCCTATGGACACCCCAATCCTGGGATGTTCCCCTCTGGAACCTCAGGAGGACCATATG  
GCGGTGCAGCTCCCGGGGGCCCCCTATGGTCAGCCACCTCCAAGTTCCTACGGTGCCCAGCAGCCT  
GGGCTTTATGGACAGGGTGGCGCCCCCTCCCAATGTGGATCCTGAGGCCTACTCCTGGTTCCAGTC  
GGTGGACTCAGATCACAGTGGCTATATCTCCATGAAGGAGCTAAAGCAGGCCCTGGTCAACTGCA  
ATTGGTCTTCATTCAATGATGAGACCTGCCTCATGATGATAAACATGTTTGACAAGACCAAGTCA  
GGCCGCATCGATGTCTACGGCTTCTCAGCCCTGTGGAAATTCATCCAGCAGTGGAAGAACCTCTT  
CCAGCAGTATGACCGGGACCGCTCGGGCTCCATTAGCTACACAGAGCTGCAGCAAGCTCTGTCCC  
AAATGGGCTACAACCTGAGCCCCCAGTTCACCCAGCTTCTGGTCTCCCGCTACTGCCCACGCTCT  
GCCAATCCTGCCATGCAGCTTGACCGCTTCATCCAGGTGTGCACCCAGCTGCAGGTGCTGACAGA  
GGCCTTCCGGGAGAAGGACACAGCTGTACAAGGCAACATCCGGCTCAGCTTCGAGGACTTCGTCA  
CCATGACAGCTTCTCGGATGCTATGACCCAACCATCTGTGGAGAGTGGAGTGCACCAGGGACCTT  
TCCTGGCTTCTTAGAGTGAGAGAAGTATGTGGACATCTCTTCTTTCTGTCCCTCTAGAAGAAC  
ATTCTCCCTTGCTTGATGCAACACTGTTCCAAAAGAGGGTGGAGAGTCCTGCATCATAGCCACCA  
AATAGTGAGGACCGGGGCTGAGGCCACACAGATAGGGGCCTGATGGAGGAGAGGATAGAAGTTGA  
ATGTCCTGATGGCCATGAGCAGTTGAGTGGCACAGCCTGGCACCAGGAGCAGGTCCTTGTAATGG  
AGTTAGTGTCCAGTCAGCTGAGCTCCACCCTGATGCCAGTGGTGAGTGTTTCATCGGCCTGTTACC  
GTTAGTACCTGTGTTCCCTCACCAGGCCATCCTGTCAAACGAGCCCATTTTCTCCAAAGTGGAAT  
CTGACCAAGCATGAGAGAGATCTGTCTATGGGACCAGTGGCTTGATTCTGCCACACCCATAAAT  
CCTTGTGTGTTAACTTCTAGCTGCCTGGGGCTGGCCCTGCTCAGACAAATCTGCTCCCTGGGCAT  
CTTTGGCCAGGCTTCTGCCCCCTGCAGCTGGGACCCCTCACTTGCTGCCATGCTCTGCTCGGCT  
TCAGTCTCCAGGAGACAGTGGTCACCTCTCCCTGCCAATACTTTTTTTAATTTGCATTTTTTTTC  
ATTTGGGGCCAAAAGTCCAGTGAAATTGTAAGCTTCAATAAAAGGATGAAACTCTGA

## **FIGURE 62**

MASYPYRQGCPGAAGQAPGAPPGSYYPGPPNSGGQYGSGLPPGGGYGGPAPGGPYGPPAGGGPYG  
HPNPGMFPSGTPGGPYGGAAPGGPYGQPPSSYGAQQPGLYGQGGAPPNVDPEAYSWFQSVDS DH  
SGYISMKELKQALVNCNWSSFNDETCLMMINMFDKTKSGRIDVYGFSALWKFIQQWKNLFQQYDR  
DRSGSISYTELQQALSQMGYNLSPQFTQLLVSRYCPRSANPAMQLDRFIQVCTQLQVLTEAFREK  
DTAVQGNIRLSFEDFVTMTASRML

**Important features of the protein:**

**Signal peptide:**

amino acids 1-19

**N-glycosylation site.**

amino acids 147-150

**Casein kinase II phosphorylation sites.**

amino acids 135-138, 150-153, 202-205, 271-274

**N-myristoylation sites.**

amino acids 9-14, 15-20, 19-24, 33-38, 34-39, 39-44, 43-48, 61-66,  
70-75, 78-83, 83-88, 87-92, 110-115

**FIGURE 63**

CAGG**ATG**CAGGGCCGCGTGGCAGGGAGCTGCGCTCCTCTGGGCCTGCTCCTGGTCTGTCTTCATC  
TCCCAGGCCTCTTTGCCCCGAGCATCGGTGTTGTGGAGGAGAAAGTTTCCCAAACTTCGGGACC  
AACTTGCCTCAGCTCGGACAACCTTCCTCCACTGGCCCCCTCTAACTCTGAACATCCGCAGCCCGC  
TCTGGACCCTAGGTCTAATGACTTGGCAAGGGTTCCTCTGAAGCTCAGCGTGCCTCCATCAGATG  
GCTTCCACCTGCAGGAGGTTCTGCAGTGCAGAGGTGGCCTCCATCGTGGGGGCTGCCTGCCATG  
GATTCCTGGCCCCCTGAGGATCCTTGGCAGATGATGGCTGCTGCGGCTGAGGACCGCCTGGGGGA  
AGCGCTGCCTGAAGAACTCTCTTACCTCTCCAGTGTGCGGCCCTCGCTCCGGGCAGTGGCCCTT  
TGCCTGGGGAGTCTTCTCCCGATGCCACAGGCCTCTCACCTGAGGCTTCACTCCTCCACCAGGAC  
TCGGAGTCCAGACGACTGCCCCGTTCTAATTCAGTGGGAGCCGGGGGAAAAATCCTTTCCCAACG  
CCCTCCCTGGTCTCTCATCCACAGGGTCTGCCTGATCACCCCTGGGGTACCCTGAATCCCAGTG  
TGTCTGGGGAGGTGGAGGCCCTGGGACTGGTTGGGGAACGAGGCCCATGCCACACCCTGAGGGA  
ATCTGGGGTATCAATAATCAACCCCCAGGTACCAGCTGGGGAAATATTAATCGGTATCCAGGAGG  
CAGCTGGGGAAATATTAATCGGTATCCAGGAGGCAGCTGGGGGAATATTAATCGGTATCCAGGAG  
GCAGCTGGGGGAATATTCATCTATACCCAGGTATCAATAACCCATTTCTCCTGGAGTTCTCCGC  
CCTCCTGGCTCTTCTTGGAACATCCCAGCTGGCTTCCCTAATCCTCCAAGCCCTAGGTTGCAGTG  
GGGC**TAG**AGCACGATAGAGGGAAACCCAACATTGGGAGTTAGAGTCCTGCTCCCGCCCCCTTGCTG  
TGTGGGCTCAATCCAGGCCCTGTAAACATGTTTCCAGCACTATCCCCACTTTTCAGTGCCTCCCC  
TGCTCATCTCCAATAAAATAAAAGCACTTATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAA

## **FIGURE 64**

MQGRVAGSCAPLGLLLVLCLHLPGLFARSIGVVEEKVSQNFGTNLPQLGQPSSTGPSNSEHPQPAL  
DPRSNDLARVPLKLSVPPSDGFPFAGGSAVQRWPPSWGLPAMDSWPPEDPWQMMAAAAEDRLGEA  
LPEELSYLSSAAALAPGSGPLPGESSPDATGLSPEASLLHQDSESRRLPRSNSLGAGGKILSQRP  
PWSLIHRVLPDHPWGTLNPSVSWGGGGPGTGWGTRPMPHPGEGIWGINNQPPGTSGWNINRYPGGS  
WGNINRYPGGSWGNINRYPGGSWGNIHLYPGINNPFPFPGVLRPPGSSWNIIPAGFPNPPSPRLQWG

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-26

#### **Casein kinase II phosphorylation sites.**

amino acids 56-59, 155-158

#### **N-myristoylation sites.**

amino acids 48-53, 220-225, 221-226, 224-229, 247-252, 258-263,  
259-264, 269-274, 270-275, 280-285, 281-286, 305-310

## **FIGURE 65**

AAGGAGAGGCCACCGGGACTTCAGTGTCTCCTCCATCCCAGGAGCGCAGTGGCCACTATGGGGTC  
TGGGCTGCCCCTTGTCTCCTCTTGACCCTCCTTGGCAGCTCACATGGAACAGGGCCGGGTATGA  
CTTTGCAACTGAAGCTGAAGGAGTCTTTTCTGACAAATTCCTCCTATGAGTCCAGCTTCCTGGAA  
TTGCTTGAAAAGCTCTGCCTCCTCCTCCATCTCCCTTCAGGGACCAGCGTCACCCTCCACCATGC  
AAGATCTCAACACCATGTTGTCTGCAACACATGACAGCCATTGAAGCCTGTGTCCTTCTTGGCCC  
GGGCTTTTGGGCCGGGGATGCAGGAGGCAGGCCCCGACCCTGTCTTTCAGCAGGCCCCCACCCTC  
CTGAGTGGCAATAAATAAAATTTCGGTATGCTG

## **FIGURE 66**

MGSGPLVLLLTLTGSSHGTGPGMTLQLKLKESFLTNSSESSFLELLEKLCLLLHLPSTSVTL  
HHARSQHHVVCNT

**Important features:**

**Signal peptide:**

amino acids 1-19

**N-glycosylation site.**

amino acids 37-41

**N-myristoylation sites.**

amino acids 15-21, 19-25, 60-66

**FIGURE 67**

ACGGACCGAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCCCGGGC  
CAGGTGCCCCGTCGCAGGTGCCCCTGGCCGGAGATGCGGTAGGAGGGGCGAGCGCGAGAAGCCCC  
TTCCTCGGCGCTGCCAACCCGCCACCCAGCCCATGGCGAACCCCGGGCTGGGGCTGCTTCTGGCG  
CTGGGCCTGCCGTTCTTGCTGGCCCGCTGGGGCCGAGCCTGGGGGCAAATACAGACCACTTCTGC  
AAATGAGAATAGCACTGTTTTGCCTTCATCCACCAGCTCCAGCTCCGATGGCAACCTGCGTCCGG  
AAGCCATCACTGCTATCATCGTGGTCTTCTCCCTCTTGGCTGCCTTGCTCCTGGCTGTGGGGCTG  
GCACTGTTGGTGCGGAAGCTTCGGGAGAAGCGGCAGACGGAGGGCACCTACCGGCCCAGTAGCGA  
GGAGCAGTTCTCCCATGCAGCCGAGGCCCGGGCCCCCTCAGGACTCCAAGGAGACGGTGCAGGGCT  
GCCTGCCCATCTAGGTCCCCTCTCCTGCATCTGTCTCCCTTCATTGCTGTGTGACCTTGGGGAAA  
GGCAGTGCCCTCTCTGGGCAGTCAGATCCACCCAGTGCTTAATAGCAGGGAAGAAGGTACTTCAA  
AGACTCTGCCCCTGAGGTCAAGAGAGGATGGGGCTATTCACTTTTATATATTTATATAAAATTAG  
TAGTGAGATGTAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 68**

MANPGLGLLLALGLPFLRLARWGRAWGQIQTTSANENSTVLPSSSTSSSSDGNLRPEAITAIIVVFS  
LLAALLLAVGLALLVRKLREKRQTEGTYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

**Important features:**

**Signal peptide:**

amino acids 1-19

**Transmembrane domain:**

amino acids 56-80

**N-glycosylation site.**

amino acids 36-40

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 86-90

**Tyrosine kinase phosphorylation site.**

amino acids 86-94

**N-myristoylation sites.**

amino acids 7-13, 26-32



[illegible]

**FIGURE 70**

MGLFRGFVFLVLCLLHQSNTSFIKLNNGFEDIVIVIDPSVPEDEKIIEQIEDMVTTASTYLFE  
 ATEKRFFFFKNVSILIPENWKENPQYKRPKHENHKKHADVIVAPPTLPGRDEPYTKQFTECGEKGEY  
 IHFTPDLLLGGKKQNEYGPPGKLFVHEWAHLRWGVFDEYNEDQPFYRAKSKKIEATRCISAGISGRN  
 RVYKCQGGSCLSRACRIDSTTKLYGKDCQFFPDQVQTEKASIMFMQSIDSVVEFCNEKTHNQEAP  
 SLQNIKCNFRSTWEVISNSEDFKNTIPMVTPPPPVFSLLKISQRIVCLVLDKSGSMGGKDRLNR  
 MNQAAKHFLQTVENGSWVGMVHFDSTATIVNKLIQIKSSDERNTLMAGLPTYPLGGTSICSGIK  
 YAFQVIGELHSQLDGSEVLLLTDGEDNTASSCIDEVKQSGAIVHFIALGRAADEAVIEMSKITGG  
 SHFYVSDEAQNNGLIDAFGALTSGNTDLSQKSLQLESKGLTLNSNAWMNDTVIIDSTVGKDTFFL  
 ITWNSLPPSISLWDPSGTIMENFTVDATSKMAYLSIPGTAKVGTWAYNLQAKANPETLTITVTSR  
 AANSSVPPITVNAKMNKDVNSFPSPMIVYAEILQGYVPVLGANVTAFIESQNGHTEVLELLDNGA  
 GADSFKNDGVYSRYFTAYTENGRYSLKVRAHGGANTARLKLRPPLNRAAYIPGWVVNGEIEANPP  
 RPEIDEDTQTTLEDFSRASGGAFVVSQVPSLPLPDQYPPSQITDLDATVHEDKIIILTWTAPGDN  
 FDVGKVQRYIIRISASILDLRDSFDDALQVNTTDLSPKEANSKESFAFKPENISEENATHIFIAI  
 KSIDKSNLTSKVSNIAQVTLFIPQANPDDIDPTPTPTPTPTPDKSHNSGVNISTLVLSVIGSVVI  
 VNFILSTTI

**Signal peptide:**

amino acids 1-21

**Putative transmembrane domains:**

amino acids 284-300, 617-633

**Leucine zipper pattern.**

amino acids 469-491, 476-498

**N-glycosylation site.**

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592,  
 628-632, 811-815, 832-836, 837-841, 852-856, 896-900

**FIGURE 71**

CTCCTTAGGTGGAAACCCTGGGAGTAGAGTACTGACAGCAAAGACCGGGAAAGACCATACGTCCCCGGGCAGGGGTGA  
CAACAGGTGTCATCTTTTTGATCTCGTGTGTGGCTGCCTTCCTATTTCAAGGAAAGACGCCAAGGTAATTTTGACCCA  
GAGGAGCAATGATGTAGCCACCTCCTAACCTTCCCTTCTTGAACCCCCAGTTATGCCAGGATTTACTAGAGAGTGTCA  
ACTCAACCAGCAAGCGGCTCCTTCGGCTTAACTTGTGGTTGGAGGAGAGAACCTTTGTGGGGCTGCGTTCTCTTAGCA  
GTGCTCAGAAGTGACTTGCCTGAGGGTGGACCAGAAGAAAGGAAAGGTCCCCTCTTGCTGTTGGCTGCACATCAGGAA  
GGCTGTGATGGGAATGAAGGTGAAAÇTTGGAGATTTCACTTCAGTCATTGCTTCTGCCTGCAAGATCATCCTTTAAA  
AGTAGAGAAGCTGCTCTGTGTGGTGGTTAACTCCAAGAGGCAGAACTCGTTCTAGAAGGAAATGGATGCAAGCAGCTC  
CGGGGGCCCCAAACGCATGCTTCTGTGGTCTAGCCCAGGGAAGCCCTTCCGTGGGGGGCCCCGGCTTTGAGGGATGCC  
ACCGGTTCTGGACGCATGGCTGATTCTGAATGATGATGGTTCCGCCGGGGCTGCTTGCCTGGATTTCCTGGGTGGTG  
GTTTTGTGGTGCTCCTCTGCTGTGCTATCTCTGCTGTACATGTTGGCCTGCACCCCAAAGGTGACGAGGAGCAG  
CTGGCACTGCCCAGGGCCAACAGCCCCACGGGGAAGGAGGGGTACCAGGCCGTCTTTCAGGAGTGGGAGGAGCAGCAC  
CGCAACTACGTGAGCAGCCTGAAGCGGCAGATCGCACAGCTCAAGGAGGAGCTGCAGGAGAGGAGTGAAGCAGCTCAGG  
AATGGGCAGTACCAAGCCAGCGATGCTGCTGGCCTGGGTCTGGACAGGAGCCCCCAGAGAAAACCCAGGCCGACCTC  
CTGGCCTTCTGCACTCGCAGGTGGACAAGGCAGAGGTGAATGCTGGCGTCAAGCTGGCCACAGAGTATGCAGCAGTG  
CCTTTTCGATAGCTTTACTCTACAGAAGGTGTACCAGCTGGAGACTGGCCTTACCCGCCACCCCGAGGAGAAGCCTGTG  
AGGAAGGACAAGCGGGATGAGTTGGTGGAAGCCATTGAATCAGCCTTGGAGACCCTGAACAATCCTGCAGAGAACAGC  
CCCAATCACCCTCCTTACACGGCCTCTGATTTTCATAGAAGGGATCTACCGAACAGAAAGGGACAAAGGGACATTGTAT  
GAGCTCACCTTCAAAGGGGACCACAAACACGAATTCAAACGGCTCATCTTATTTTCGACCATTTCAGCCCCATCATGAAA  
GTGAAAATGAAAAGCTCAACATGGCCAACACGCTTATCAATGTTATCGTGCCTCTAGCAAAAAGGGTGGACAAGTTC  
CGGCAGTTCATGCAGAATTTTCAGGGAGATGTGCATTGAGCAGGATGGGAGAGTCCATCTCACTGTTGTTTACTTTGGG  
AAAGAAGAAATAAATGAAGTCAAAGGAATACTTGAAAACACTTCCAAAGCTGCCAACTTCAGGAACCTTTACCTTCATC  
CAGCTGAATGGAGAATTTTCTCGGGGAAGGGACTTGATGTTGGAGCCCGCTTCTGGAAGGGAAGCAACGTCTTCTC  
TTTTTCTGTGATGTGGACATCTACTTCACATCTGAATTCCTCAATACGTGTAGGCTGAATACACAGCCAGGGAAGAAG  
GTATTTTATCCAGTTCCTTTTCAGTCAGTACAATCCTGGCATAATATACGGCCACCATGATGCAGTCCCTCCCTTGAA  
CAGCAGCTGGTCATAAAGAAGGAACTGGATTTTGGAGAGACTTTGGATTGGGATGACGTGTCAAGTATCGGTGAGC  
TTCATCAATATAGGTGGGTTTGATCTGGACATCAAAGGCTGGGGCGGAGAGGATGTGCACCTTTATCGCAAGTATCTC  
CACAGCAACCTCATAGTGGTACGGACGCCTGTGCGAGGACTCTTCCACCTCTGGCATGAGAAGCGCTGCATGGACGAG  
CTGACCCCGGAGCAGTACAAGATGTGCATGCAGTCCAAGGCCATGAACGAGGCATCCACGGCCAGCTGGGCATGCTG  
GTGTTTCAGGCACGAGATAGAGGCTCACCTTCGCAACAGAAACAGAAGACAAGTAGCAAAAAACATGAAGTCCCAGA  
GAAGGATTGTGGGAGACACTTTTTCTTTCTTTTGCAATTACTGAAAGTGGCTGCAACAGAGAAAAGACTTCCATAAA  
GGACGACAAAAGAATTGGACTGATGGGTGAGAGATGAGAAAGCCTCCGATTTCTCTCTGTTGGGCTTTTTACAACAGA  
AATCAAAATCTCCGCTTTGCCTGCAAAAGTAACCCAGTTGCACCCTGTGAAGTGTCTGACAAAGGCAGAATGCTTGTG  
AGATTATAAGCCTAATGGTGTGGAGGTTTTGATGGTGTTTACAATACACTGAGACCTGTTGTTTTGTGTGCTCATTGA  
AATATTCATGATTTAAGAGCAGTTTTGTAAAAAATTCATTAGCATGAAAGGCAAGCATATTTCTCCTCATATGAATGA  
GCCTATCAGCAGGGCTCTAGTTTTCTAGGAATGCTAAAATATCAGAAGGCAGGAGAGGAGATAGGCTTATTATGATACT  
AGTGAGTACATTAAGTAAAATAAAATGGACCAGAAAAGAAAAGAAACCATAAATATCGTGTCTATTTTTCCCCAAGAT  
TAACCAAAAATAATCTGCTTATCTTTTTGGTTGTCTTTTAACTGTCTCCGTTTTTTTTCTTTTATTTAAAAATGCACT  
TTTTTTCCCTTGTGAGTTATAGTCTGCTTATTTAATTACCCTTTGCAAGCCTTACAAGAGAGCACAAGTTGGCCTAC  
ATTTTTATATTTTTTAAGAAGATACTTTGAGATGCATTATGAGAACTTTTCAGTTCAAAGCATCAAATTGATGCCATAT  
CCAAGGACATGCCAAATGCTGATTCTGTGAGGCACTGAATGTGAGGCATTGAGACATAGGGAAGGAATGGTTTTGTACT  
AATACAGACGTACAGATACTTTCTCTGAAGAGTATTTTCGAAGAGGAGCAACTGAACACTGGAGGAAAAGAAAATGAC  
ACTTTCTGCTTTACAGAAAAGGAACTCATTAGCATGGTGATATCGTGATGTACCTAAAAGTCAGAAACCACATTTT  
CTCCTCAGAAGTAGGGACCGCTTTTCTACCTGTTTAAATAAACCAAGTATACCGTGTGAACCAACAATCTCTTTTC  
AAAACAGGGTGTCTCCTCGCTTCTGGCTTCCATAAGAAGAAATGGAGAAAAATATATATATATATATATATATTGT  
GAAAGATCAATCCATCTGCCAGAATCTAGTGGGATGGAAGTTTTTGTACATGTTATCCACCCAGGCCAGGTGGAAG  
TAACTGAATTTATTTTTAAATTAAGCAGTTCTACTCAATCACCAGATGCTTCTGAAAATTGCATTTTATTACCATTT  
CAAATATTTTTTAAAAATAAATACAGTTAATAGAGTGGTTTTCTTCATTCATGTGAAAATTATTAGCCAGCACCAG  
ATGCATGAGCTAATTATCTCTTTGAGTCCTTGCTTCTGTTTGTCTCACAGTAAACTCATTGTTTTAAAGCTTCAAGAAC  
ATTCAGCTGTTGGTGTGTTAAAAAATGCATTGTATTGATTTGTACTGGTAGTTTATGAAATTTAATTAAAAACACAGG  
CCATGAATGGAAGGTGGTATTGCACAGCTAATAAAATATGATTTGTGGATATGAA

**FIGURE 72**

MMVRRGLLAWISRVVLLVLLCCAISVLYMLACTPKGDEEQLALPRANSPTGKEGYQAVLQEW  
EQHRNYVSSSLKRQIAQLKEELQERSEQLRNGQYQASDAAGLGLDRSPPEKTQADLLAFLHSQVDK  
AEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPEEKPV RKDKRDELVEAIESALET LNNPA  
ENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSPIMKV KNEKLNMAN  
TLINVIVPLAKRVDKFRQFMQNFREMCIEQDGRVH LTVVYFGKEEINEVKGILENTSKAANFRNF  
TFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNTCRLNTQPGKKVFYPVLFSQY  
NPGIIYGHHDVAPPLEQQLVIKKETGFWRDFGFGMTCQYRSDFINIGGFDLDIKGWGGEDVHLYR  
KYLHSNLIVVRTPV RGLFHLWHEKRCMDELTP EQYKMCMQSKAMNEASHGQLGMLVFRHEIEAHL  
RKQKQKTSSKKT

**Important features:****Signal peptide:**

amino acids 1-27

**N-glycosylation sites.**

amino acids 315-319, 324-328

**N-myristoylation sites.**

amino acids 96-102, 136-142, 212-218, 311-317, 339-345, 393-399

**Amidation site.**

amino acids 377-381

**FIGURE 73**

GAGACTGCAGAGGGAGATAAAGAGAGAGGGCAAAGAGGCAGCAAGAGATTTGTCCTGGGGATCCA  
GAAACCCATGATACCCTACTGAACACCGAATCCCCTGGAAGCCCACAGAGACAGAGACAGCAAGA  
GAAGCAGAGATAAATACACTCACGCCAGGAGCTCGCTCGCTCTCTCTCTCTCTCACTCCTC  
CCTCCCTCTCTCTCTGCCTGTCCTAGTCCTCTAGTCCTCAAATTCCCAGTCCCCTGCACCCCTTC  
CTGGGACACT**ATG**TTGTTCTCCGCCCTCCTGCTGGAGGTGATTTGGATCCTGGCTGCAGATGGGG  
GTCAACACTGGACGTATGAGGGCCCCACATGGTCAGGACCATTGGCCAGCCTCTTACCCTGAGTGT  
GGAAACAATGCCCAGTCGCCCATCGATATTCAGACAGACAGTGTGACATTTGACCCTGATTTGCC  
TGCTCTGCAGCCCCACGGATATGACCAGCCTGGCACCAGCCTTTGGACCTGCACAACAATGGCC  
ACACAGTGCAACTCTCTCTGCCCTCTACCCTGTATCTGGGTGGACTTCCCCGAAAATATGTAGCT  
GCCAGCTCCACCTGCACTGGGGTCAGAAAGGATCCCCAGGGGGGTGAGAACACCAGATCAACAG  
TGAAGCCACATTTGCAGAGCTCCACATTGTACATTATGACTCTGATTCCTATGACAGCTTGAGTG  
AGGCTGCTGAGAGGCCTCAGGGCCTGGCTGTCTGGGCATCCTAATTGAGGTGGGTGAGACTAAG  
AATATAGCTTATGAACACATTCTGAGTCACTTGCATGAAGTCAGGCATAAAGATCAGAAGACCTC  
AGTGCCTCCCTTCAACCTAAGAGAGCTGCTCCCCAAACAGCTGGGGCAGTACTTCCGCTACAATG  
GCTCGCTCACAACCTCCCCCTTGCTACCAGAGTGTGCTCTGGACAGTTTTTTTATAGAAGGTCCCAG  
ATTTCAATGGAACAGCTGGAAAAGCTTCAGGGGACATTGTTCTCCACAGAAGAGGAGCCCTCTAA  
GCTTCTGGTACAGAACTACCGAGCCCTTCAGCCTCTCAATCAGCGCATGGTCTTTGCTTCTTTCA  
TCCAAGCAGGATCCTCGTATACCACAGGTGAAATGCTGAGTCTAGGTGTAGGAATCTTGGTTGGC  
TGTCTCTGCCTTCTCCTGGCTGTTTATTTCAATTGCTAGAAAGATTCCGAAGAAGAGGCTGGAAAA  
CCGAAAGAGTGTGGTCTTCACCTCAGCACAAGCCACGACTGAGGCAT**TAA**ATTCCTTCTCAGATAC  
CATGGATGTGGATGACTTCCCTTCATGCCTATCAGGAAGCCTCTAAAATGGGGTGTAGGATCTGG  
CCAGAAACACTGTAGGAGTAGTAAGCAGATGTCCTCCTTCCCCTGGACATCTCTTAGAGAGGAAT  
GGACCCAGGCTGTCATTCCAGGAAGAACTGCAGAGCCTTCAGCCTCTCCAAACATGTAGGAGGAA  
ATGAGGAAATCGCTGTGTTGTTAATGCAGAGANCAAACTCTGTTTAGTTGCAGGGGAAGTTTGGG  
ATATACCCCAAAGTCCTCTACCCCTCACTTTTATGGCCCTTTCCCTAGATATACTGCGGGATCT  
CTCCTTAGGATAAAGAGTTGCTGTTGAAGTTGTATATTTTTGATCAATATATTTGGAAATTAAAG  
TTTCTGACTTT

## **FIGURE 74**

MLFSALLLEVIWILAADGGQHWTYEGPHGQDHWPA SYPECGNNAQSPIDIQTDSVTFD PDLPALQ  
PHGYDQPGTEPLDLHNNGHTVQLSLPSTLYLGGLPRKYVAAQLHLHWGQKGSPGGSEHQINSEAT  
FAELHIVHYDSYDSLSEAAERPQGLAVLGILIEVGETKNIAYEHILSHLHEVRHKDQKTSVPP  
FNLRELLPKQLGQYFRYNGSLTTPPCYQSVLWTVFYRRSQISMEQLEKLQGTLFSTEEEPSKLLV  
QNYRALQPLNQRMVFASFIQAGSSYTTGEMLSLGVGILVGCLCLLLAVYFIARKIRKKRLENRKS  
VVFTSAQATTEA

**Important features of the protein:**

**Signal peptide:**

amino acids 1-15

**Transmembrane domain:**

amino acids 291-310

**N-glycosylation site.**

amino acids 213-216

**Eukaryotic-type carbonic anhydrases proteins**

amino acids 197-245, 104-140, 22-69

**FIGURE 75**

TGCCGCTGCCGCCGCTGCTGCTGTTGCTCCTGGCGGCGCCTTGGGGACGGGCAGTTCCCTGTGTC  
TCTGGTGGTTTGCCTAAACCTGCAAACATCACCTTCTTATCCATCAACATGAAGA**ATG**TCCTACA  
ATGGACTCCACCAGAGGGTCTTCAAGGAGTTAAAGTTACTTACACTGTGCAGTATTTTCATCACAA  
ATTGGCCCACCAGAGGTGGCACTGACTACAGATGAGAAGTCCATTTCTGTTGTCCTGACAGCTCC  
AGAGAAGTGGAAGAGAAATCCAGAAGACCTTCTGTTTCCATGCAACAAATATACTCCAATCTGA  
AGTATAACGTGTCTGTGTTGAATACTAAATCAAACAGAACGTGGTCCCAGTGTGTGACCAACCAC  
ACGCTGGTGCTCACCTGGCTGGAGCCGAACACTCTTTACTGCGTACACGTGGAGTCCTTCGTCCC  
AGGGCCCCCTCGCCGTGCTCAGCCTTCTGAGAAGCAGTGTGCCAGGACTTTGAAAGATCAATCAT  
CAGAGTTCAAGGCTAAAATCATCTTCTGGTATGTTTTGCCCATATCTATTACCGTGTTTCTTTTT  
TCTGTGATGGGCTATTCCATCTACCGATATATCCACGTGGCAAAGAGAAACACCCAGCAAATTT  
GATTTTGATTTATGGAAATGAATTTGACAAAAGATTCTTTGTGCCTGCTGAAAAAATCGTGATTA  
ACTTTATCACCCCTCAATATCTCGGATGATTCTAAAATTTCTCATCAGGATATGAGTTTACTGGGA  
AAAAGCAGTGATGTATCCAGCCTTAATGATCCTCAGCCAGCGGGAACCTGAGGCCCCCTCAGGA  
GGAAGAGGAGGTGAAACATTTAGGGTATGCTTCGCATTTGATGGAAATTTTTTGACTCTGAAG  
AAAACACGGAAGGTACTTCTCTCACCCAGCAAGAGTCCCTCAGCAGAACAAATACCCCCGGATAAA  
ACAGTCATTGAATATGAATATGATGTCAGAACCACTGACATTTGTGCGGGGCCTGAAGAGCAGGA  
GCTCAGTTTGCAGGAGGAGGTGTCCACACAAGGAACATTATTGGAGTCGCAGGCAGCGTTGGCAG  
TCTTGGGCCCCGAAACGTTACAGTACTCATAACCCCTCAGCTCCAAGACTTAGACCCCTGGCG  
CAGGAGCACACAGACTCGGAGGAGGGGCCGAGGAAGAGCCATCGACGACCCTGGTCGACTGGGA  
TCCCCAACTGGCAGGCTGTGTATTCTTCGCTGTCCAGCTTCGACCAGGATTAGAGGGCTGCG  
AGCCTTCTGAGGGGGATGGGCTCGGAGAGGAGGGTCTTCTATCTAGACTCTATGAGGAGCCGGCT  
CCAGACAGGCCACCAGGAGAAAATGAAACCTATCTCATGCAATTCATGGAGGAATGGGGGTTATA  
TGTGCAGATGGAAAAC**TGA**TGCCAACACTTCCTTTTGCCTTTTGTTCCTGTGCAAACAAGTGAG  
TCACCCCTTTGATCCCAGCCATAAAGTACCTGGGATGAAAGAAGTTTTTCCAGTTTGTGTCAGTGT  
CTGTGAGAATTACTTATTTCTTTTCTCTATTCTCATAGCACGTGTGTGATTGGTTCATGCATGTA  
GGTCTCTTAACAATGATGGTGGGCCTCTGGAGTCCAGGGGCTGGCCGGTTGTTCTATGCAGAGAA  
AGCAGTCAATAAATGTTTGCCAGACTGGGTGCAGAATTTATTAGGTGGGTGT

## **FIGURE 76**

MSYNGLHQRVFKELKLLTLCSSISQIGPPEVALTTDEKSISVVLTAPEKWKRNPEDLPVSMQQIY  
SNLKYNVSVLNTKSNRTWSQCVTNHTLVLTWLEPNTLYCVHVESFVPGPPRAQPSEKQCARTLK  
DQSSEFKAKIIFWYVLPISITVFLFSVMGYSIYRYIHVGKEKHPANLILYGNFEDKRFFVPAEK  
IVINFITLNISSDDSKISHQDMSLLGKSSDVSSLNDPQPSGNLRPPQEEEEVKHLGYASHLMEIFC  
DSEENTEGTSLTQQESLSRTIPPDKTVIEYEDVRTTDICAGPEEQELSLQEEVSTQGTLLLESQA  
ALAVLGPQTLQYSYTPQLQDLPLAQEHTDSEEGPEEEPSTTLVDWDPQTGRLCIPSLSSFDQDS  
EGCEPSEGDGLGEEGLLSRLYEPPAPDRPPGENETYLMQFMEEWGLYVQMEN

**Important features:**

**Signal peptide:**

amino acids 1-28

**Transmembrane domain:**

amino acids 140-163

**N-glycosylation sites.**

amino acids 71-74, 80-83, 89-92, 204-207, 423-426



**FIGURE 77**

GAGGAGCGGGCCGAGGACTCCAGCGTGCCCAAGTCTGGCATCCTGCACTTGCTGCCCTCTGACAC  
CTGGGAAG**ATG**GGCCGGCCCGTGGACCTTCACCCCTTCTCTGTGGTTTGCTGGCAGCCACCTTGATC  
CAAGCCACCCTCAGTCCCCTGACAGTTCTCATCCTCGGCCCAAAGTCATCAAAGAAAAGCTGAC  
ACAGGAGCTGAAGGACCACAACGCCACCAGCATCCTGCAGCAGCTGCCGCTGCTCAGTGCCATGC  
GGGAAAAGCCAGCCGGAGGCATCCCTGTGCTGGGCAGCCTGGTGAACACCGTCCTGAAGCACATC  
ATCTGGCTGAAGGTCATCACAGCTAACATCCTCCAGCTGCAGGTGAAGCCCTCGGCCAATGACCA  
GGAGCTGCTAGTCAAGATCCCCCTGGACATGGTGGCTGGATTCAACACGCCCTGGTCAAGACCA  
TCGTGGAGTTCCACATGACGACTGAGGCCCAAGCCACCATCCGCATGGACACCAGTGCAAGTGGC  
CCCACCCGCCTGGTCCTCAGTGACTGTGCCACCAGCCATGGGAGCCTGCGCATCCAACTGCTGTA  
TAAGCTCTCCTTCCCTGGTGAACGCCTTAGCTAAGCAGGTCATGAACCTCCTAGTGCCATCCCTGC  
CCAATCTAGTGAAAAACCAGCTGTGTCCCGTGATCGAGGCTTCCTTCAATGGCATGTATGCAGAC  
CTCCTGCAGCTGGTGAAGGTGCCCATTTCCTCAGCATTGACCGTCTGGAGTTTGACCTTCTGTA  
TCCTGCCATCAAGGGTGACACCATTACAGCTCTACCTGGGGGCCAAGTTGTTGGACTCACAGGGAA  
AGGTGACCAAGTGGTTCAATAACTCTGCAGCTTCCCTGACAATGCCACCCTGGACAACATCCCG  
TTCAGCCTCATCGTGAGTCAGGACGTGGTGAAGCTGCAGTGGCTGCTGTGCTCTCTCCAGAAGA  
ATTCATGGTCCTGTTGGACTCTGTGCTTCCTGAGAGTGCCCATCGGCTGAAGTCAAGCATCGGGC  
TGATCAATGAAAAGGCTGCAGATAAGCTGGGATCTACCCAGATCGTGAAGATCCTAACTCAGGAC  
ACTCCCGAGTTTTTTATAGACCAAGGCCATGCCAAGGTGGCCCAACTGATCGTGCTGGAAGTGTT  
TCCCTCCAGTGAAGCCCTCCGCCCTTTGTTACCCCTGGGCATCGAAGCCAGCTCGGAAGCTCAGT  
TTTACACCAAAGGTGACCAACTTATACTCAACTTGAATAACATCAGCTCTGATCGGATCCAGCTG  
ATGAACTCTGGGATTGGCTGGTTCCAACCTGATGTTCTGAAAAACATCATCACTGAGATCATCCA  
CTCCATCCTGCTGCCGAACCAGAATGGCAAATTAAGATCTGGGGTCCCAGTGTCATTGGTGAAGG  
CCTTGGGATTGAGGCAGCTGAGTCCTCACTGACCAAGGATGCCCTTGCTGCTTACTCCAGCCTCC  
TTGTGGAAACCCAGCTCTCCTGTCTCCAG**TGA**AGACTTGGATGGCAGCCATCAGGGAAGGCTGG  
GTCCAGCTGGGAGTATGGGTGTGAGCTCTATAGACCATCCCTCTCTGCAATCAATAAACACTTG  
CCTGTGAAAAA

## **FIGURE 78**

MAGPWTFTLLCGLLAATLIQATLSPTAVLILGPKVIEKLTQELKDHNATSILQQLPLLSAMREK  
PAGGIPVLGSLVNTVLKHIIWLKVITANILQLQVKPSANDQELLVKIPLDMVAGFNTPLVKTIVE  
FHMTTEAQATIRMDTSASGPTRLVLSDCATSHGSLRIQLLYKLSFLVNALAKQVMNLLVPSLPNL  
VKNQLCPVIEASFNGMYADLLQLVKVPISLSIDRLEFDLLYPAIKGDITIQLYLGAKLDSQGKVT  
KWFNNSAASLTMPITLDNIPFSLIVSQDVVKAABAVALSPPEEFMVLLDSVLPESAHLKSSIGLIN  
EKAADKLGSTQIVKILTQDTPEFFIDQGHAKVAQLIVLEVFPSSALRPLFTLGIEASSEAQFYT  
KGDQLILNLNNISSDRIQLMNSGIGWFQPDVLKNIITEIIHSILLPNQNGKLRSQVPSLVKALG  
FEAAESSLTKDALLVTPASLWKPSPPVSQ

**Important features of the protein:**

**Signal peptide:**

amino acids 1-21

**N-glycosylation sites.**

amino acids 48-51, 264-267, 401-404

**Glycosaminoglycan attachment site.**

amino acids 412-415

**LBP / BPI / CETP family proteins.**

amino acids 407-457

**FIGURE 79**

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAGAGC  
TTCAGCCTGAAGACAAGGGAGCAGTCCCTGAAGACGCTTCTACTGAGAGGTCTGCCATGGCCTCT  
CTTGGCCTCCAACTTGTGGGCTACATCCTAGGCCTTCTGGGGCTTTTGGGCACACTGGTTGCCAT  
GCTGCTCCCCAGCTGGAAAACAAGTTCTTATGTCCGTGCCAGCATTGTGACAGCAGTTGGCTTCT  
CCAAGGGCCTCTGGATGGAATGTGCCACACACAGCACAGGCATCACCCAGTGTGACATCTATAGC  
ACCCTTCTGGGCCTGCCCCGCTGACATCCAGGCTGCCCAGGCCATGATGGTGACATCCAGTGCAAT  
CTCCTCCCTGGCCTGCATTATCTCTGTGGTGGGCATGAGATGCACAGTCTTCTGCCAGGAATCCC  
GAGCCAAAGACAGAGTGGCGGTAGCAGGTGGAGTCTTTTTCATCCTTGGAGGCCTCCTGGGATTC  
ATTCCTGTTGCCTGGAATCTTCATGGGATCCTACGGGACTTCTACTCACCACTGGTGCCTGACAG  
CATGAAATTTGAGATTGGAGAGGCTCTTTACTTGGGCATTATTTCTTCCCTGTTCTCCCTGATAG  
CTGGAATCATCCTCTGCTTTTCCTGCTCATCCCAGAGAAATCGCTCCAACTACTACGATGCCTAC  
CAAGCCCAACCTCTTGCCACAAGGAGCTCTCCAAGGCCTGGTCAACCTCCCAAAGTCAAGAGTGA  
GTTCAATTCCTACAGCCTGACAGGGTATGTGTGAAGAACCAGGGGGCCAGAGCTGGGGGGTGGCTG  
GGTCTGTGAAAAACAGTGGACAGCACCCCGAGGGCCACAGGTGAGGGACACTACCACTGGATCGT  
GTCAGAAGGTGCTGCTGAGGATAGACTGACTTTGGCCATTGGATTGAGCAAAGGCAGAAATGGGG  
GCTAGTGTAACAGCATGCAGGTTGAATTGCCAAGGATGCTCGCCATGCCAGCCTTTCTGTTTTCC  
TCACCTTGCTGCTCCCCTGCCCTAAGTCCCCAACCCCTCAACTTGAAACCCCATTCCTTAAGCCA  
GGACTCAGAGGATCCCTTTGCCCTCTGGTTTACCTGGGACTCCATCCCCAAACCCACTAATCACA  
TCCCACTGACTGACCCTCTGTGATCAAAGACCCCTCTCTCTGGCTGAGGTTGGCTCTTAGCTCATT  
GCTGGGGATGGGAAGGAGAAGCAGTGGCTTTTGTGGGCATTGCTCTAACCTACTTCTCAAGCTTC  
CCTCCAAAGAACTGATTGGCCCTGGAACCTCCATCCCCTCTTGTTATGACTCCACAGTGTCCA  
GACTAATTTGTGCATGAACTGAAATAAAACCATCCTACGGTATCCAGGGAACAGAAAGCAGGATG  
CAGGATGGGAGGACAGGAAGGCAGCCTGGGACATTTAAAAAATA

## **FIGURE 80**

MASLGLQLVGYYLGLLGLLGLTLVAMLLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTGITQCD  
IYSTLLGLPADIQAAQAMMVTSSAIISSLACIISVVGMRCTVFCQESRAKDRVAVAGGVFFILGGL  
LGFIPVAWNLHGILRDFYSPLVPDSMKFEIGEALYLGIISSLFSLIAGIILCFSCSSQRNRSNYY  
DAYQAQPLATRSSPRPGQPPKVKSEFNSYSLTGYV

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-24

#### **Transmembrane domains:**

amino acids 82-102, 117-140, 163-182

#### **N-glycosylation site.**

amino acids 190-193

#### **PMP-22 / EMP / MP20 family proteins.**

amino acids 46-59

**FIGURE 81**

CCCACGCGTCCGCGCCTCTCCCTTCTGCTGGACCTTCCTTCGTCTCTCCATCTCTCCCTCCTTTC  
CCCGCGTTCTCTTTCCACCTTTCTCTTCTTCCCACCTTAGACCTCCCTTCCTGCCCTCCTTTCCT  
GCCCACCGCTGCTTCCTGGCCCTTCTCCGACCCCGCTCTAGCAGCAGACCTCCTGGGGTCTGTGG  
GTTGATCTGTGGCCCCCTGTGCCTCCGTGTCTTTTTCGTCTCCCTTCCTCCCGACTCCGCTCCCGG  
ACCAGCGGCCTGACCCTGGGGAAAGG**ATG**GTTCCCGAGGTGAGGGTCCTCTCCTCCTTGCTGGGA  
CTCGCGCTGCTCTGGTTCCCCCTGGACTCCCACGCTCGAGCCCGCCAGACATGTTCTGCCTTTT  
CCATGGGAAGAGATACTCCCCGGCGAGAGCTGGCACCCCTACTTGAGGCCACAAGGCCTGATGT  
ACTGCCTGCGCTGTACCTGCTCAGAGGGCGCCCATGTGAGTTGTTACCGCCTCCACTGTCCGCCT  
GTCCACTGCCCCCAGCCTGTGACGGAGCCACAGCAATGCTGTCCCAAGTGTGTGGAACCTCACAC  
TCCCTCTGGACTCCGGGCCCCACCAAAGTCCTGCCAGCACAAACGGGACCATGTACCAACACGGAG  
AGATCTTCAGTGCCCATGAGCTGTTCCCCTCCCGCCTGCCCAACCAGTGTGTCCTCTGCAGCTGC  
ACAGAGGGCCAGATCTACTGCGGCCTCACAACCTGCCCCGAACCAGGCTGCCCAGCACCCCTCCC  
ACTGCCAGACTCCTGCTGCCAAGCCTGCAAAGATGAGGCAAGTGAGCAATCGGATGAAGAGGACA  
GTGTGCAGTCGCTCCATGGGGTGAGACATCCTCAGGATCCATGTTCCAGTGATGCTGGGAGAAAG  
AGAGGGCCCGGGCACCCAGCCCCACTGGCCTCAGCGCCCCTCTGAGCTTCATCCCTCGCCACTT  
CAGACCCAAGGGAGCAGGCAGCACAACTGTCAAGATCGTCCTGAAGGAGAAACATAAGAAAGCCT  
GTGTGCATGGCGGGAAGACGTACTCCACGGGGAGGTGTGGCACCCGGCCTTCCGTGCCTTCGGC  
CCCTTGCCCTGCATCCTATGCACCTGTGAGGATGGCCGCCAGGACTGCCAGCGTGTGACCTGTCC  
CACCGAGTACCCCTGCCGTACCCCGAGAAAGTGGCTGGGAAGTGTGCAAGATTTGCCCAGAGG  
ACAAAGCAGACCCTGGCCACAGTGAGATCAGTTCTACCAGGTGTCCCAAGGCACCGGGCCGGGTC  
CTCGTCCACACATCGGTATCCCCAAGCCCAGACAACCTGCGTCGCTTTGCCCTGGAACACGAGGC  
CTCGGACTTGGTGGAGATCTACCTCTGGAAGCTGGTAAAAGATGAGGAACTGAGGCTCAGAGAG  
GTGAAGTACCTGGCCCAAGGCCACACAGCCAGAATCTTCCACTTGACTCAGATCAAGAAAGTCAG  
GAAGCAAGACTTCAGAAAGAGGCACAGCACTTCCGACTGCTCGCTGGCCCCCACGAAGGTCACT  
GGAACGTCTTCCTAGCCCAGACCCTGGAGCTGAAGGTCACGGCCAGTCCAGACAAAGTGACCAAG  
ACATAACAAAGACC**TAA**CAGTTGCAGATATGAGCTGTATAATTGTTGTTATTATATATTAATAAA  
TAAGAAGTTGCATTACCCTCAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 82**

MVPEVRVLSSLLGLALLWFPLDSHARARPD MFCLFHGKRYSPGESWHPYLEPQGLMYCLRTCSE  
GAHVSCYRLHCPPVHCPQPVT EPQQCCPKC VEPHTPSGLRAPPKSCQHNGTMYQHGEIFSAHELF  
PSRLPNQCVLCSCTEGQIYCGLTTCPEPGCPAPLPLPDSCCQACKDEASEQSDEEDSVQSLHGVR  
HPQDPCSSDAGRKRGP GTPAPTGLSAPLSFIPRHFRPKGAGSTTVKIVLKEKHKKACVHGGKTY S  
HGEVWHPAFRAFGPLPCILCTCEDGRQDCQRVTCPT EYPCRHP EKVAGKCKICPEDKADPGHSE  
ISSTRCPKAPGRVLVHTSVSPSPDNLRRFALEHEASDLVEIYLWKLVKDEETEAQRGEVPGPRPH  
SQNLPLDSDQESQEARLPERGTALPTARWP PRRSLERLPSPDPGAEGHGQSRQSDQDITKT

**Signal peptide:**

amino acids 1-25

**FIGURE 83**

GACAGCTGTGTCTCGATGGAGTAGACTCTCAGAACAGCGCAGTTTGGCCCTCCGCTCACGCAGAGC  
CTCTCCGTGGCTTCCGCACCTTGAGCATTAGGCCAGTTCTCCTCTTCTCTCTAATCCATCCGTCA  
CCTCTCCTGTCATCCGTTTCCATGCCGTGAGGTCCATTACAGAACACATCCATGGCTCTCATGC  
TCAGTTTGTTTCTGAGTCTCCTCAAGCTGGGATCAGGGCAGTGGCAGGTGTTTGGGCCAGACAAG  
CCTGTCCAGGCCTTGGTGGGGGAGGACGCAGCATTCTCCTGTTTCCTGTCTCCTAAGACCAATGC  
AGAGGCCATGGAAGTGCAGTTCTTCAGGGGCCAGTTCTCTAGCGTGGTCCACCTCTACAGGGACG  
GGAAGGACCAGCCATTTATGCAGATGCCACAGTATCAAGGCAGGACAAAACCTGGTGAAGGATTCT  
ATTGCGGAGGGGCGCATCTCTCTGAGGCTGGAAAACATTACTGTGTTGGATGCTGGCCTCTATGG  
GTGCAGGATTAGTTCCCAGTCTTACTACCAGAAGGCCATCTGGGAGCTACAGGTGTCAGCACTGG  
GCTCAGTTCCTCTCATTTCATCACGGGATATGTTGATAGAGACATCCAGCTACTCTGTCTAGTCC  
TCGGGCTGGTTCCCCCGGCCACAGCGAAGTGGAAGGTCCACAAGGACAGGATTTGTCCACAGA  
CTCCAGGACAAACAGAGACATGCATGGCCTGTTTGATGTGGAGATCTCTCTGACCGTCCAAGAGA  
ACGCCGGGAGCATATCCTGTTCCATGCGGCATGCTCATCTGAGCCGAGAGGTGGAATCCAGGGTA  
CAGATAGGAGATACCTTTTTTCGAGCCTATATCGTGGCACCTGGCTACCAAAGTACTGGGAATACT  
CTGCTGTGGCCTATTTTTTGGCATTGTTGGACTGAAGATTTTCTTCTCCAAATTCAGTGGAAAA  
TCCAGGCGGAACTGGACTGGAGAAGAAAGCACGGACAGGCAGAATTGAGAGACGCCCGGAAACAC  
GCAGTGGAGGTGACTCTGGATCCAGAGACGGCTCACCCGAAGCTCTGCGTTTCTGATCTGAAAAC  
TGTAACCCATAGAAAAGCTCCCCAGGAGGTGCCTCACTCTGAGAAGAGATTTACAAGGAAGAGTG  
TGGTGGCTTCTCAGAGTTTCCAAGCAGGGAAACATTACTGGGAGGTGGACGGAGGACACAATAAA  
AGGTGGCGCGTGGGAGTGTGCCGGGATGATGTGGACAGGAGGAAGGAGTACGTGACTTTGTCTCC  
CGATCATGGGTACTGGGTCCTCAGACTGAATGGAGAACATTTGTATTTACATTAAATCCCCGTT  
TTATCAGCGTCTTCCCCAGGACCCACCTACAAAAATAGGGGTCTTCTGGACTATGAGTGTGGG  
ACCATCTCCTTCTTCAACATAAATGACCAGTCCCTTATTTATACCCTGACATGTCGGTTTGAAGG  
CTTATTGAGGCCCTACATTGAGTATCCGTCTATAATGAGCAAAATGGAATCCCATAGTCATCT  
GCCAGTCACCCAGGAATCAGAGAAAGAGGCCTCTTGGCAAAGGGCCTCTGCAATCCCAGAGACA  
AGCAACAGTGAGTCCTCCTCACAGGCAACCACGCCCTTCCTCCCCAGGGGTGAAATGTAGGATGA  
ATCACATCCACATTCTTCTTTAGGGATATTAAGGTCTCTCTCCAGATCCAAAGTCCCGCAGCA  
GCCGGCCAAGGTGGCTTCCAGATGAAGGGGGACTGGCCTGTCCACATGGGAGTCAGGTGTCATGG  
CTGCCCTGAGCTGGGAGGGAAGAAGGCTGACATTACATTTAGTTTGCTCTCACTCCATCTGGCTA  
AGTGATCTTGAAATACCACCTCTCAGGTGAAGAACCGTCAGGAATTCCCATCTCACAGGCTGTGG  
TGTAGATTAAGTAGACAAGGAATGTGAATAATGCTTAGATCTTATTGATGACAGAGTGTATCCTA  
ATGGTTTGTTTATTATATTACACTTTCAGTAAAAAA

## **FIGURE 84**

MALMLSLVLSLLKLGSGQWQVFGPDKPVQALVGEDAAFSCFLSPKTNAEAMEVRFFRGQFSSVH  
LYRDGKDQPFMQMPQYQGR TKLVKDSIAEGRISLRLENITVLDAGLYGCRISSQSYQKAIWELQ  
VSALGSVPLISITGYVDRDIQLLCQSSGWFP RP TAKWKGPQGQDLSTDSRTNRDMHGLFDVEISL  
TVQENAGSISCSMRHAHLSREVESRVQIGDTFFEPISWHLATKVLGILCCGLFFGIVGLKIFFSK  
FQWKIQAE LDWRRKHGQAE LR D ARKHAVEVTLD PETAH PKLCVSDLKTVTHR KAPQEVP HSEKRF  
TRKSVVASQSFQAGKHYWEVDGGHNKRWRVGVC RDDVD RRKEYVTLS PDHGYWVLRLNGEHL YFT  
LNPRFISVFPRT PPTKIGVFLDYECGTISFFNINDQSLIYT LT CRFEGLLRPYIEYPSYNEQNGT  
PIVICPVTQESEKEASWQRASAI PETSNSESSSQATT PFLPRGEM

**Signal peptide:**

amino acids 1-17

**Transmembrane domain:**

amino acids 239-255



**FIGURE 85**

AACAGACGTTCCCTCGCGGCCCTGGCACCTCTAACCCAGACATGCTGCTGCTGCTGCTGCCCCCT  
GCTCTGGGGGAGGGAGAGGGCGGAAGGACAGACAAGTAACTGCTGACGATGCAGAGTTCCGTGA  
CGGTGCAGGAAGGCCTGTGTGTCCATGTGCCCTGCTCCTTCTCCTACCCCTCGCATGGCTGGATT  
TACCCTGGCCCAGTAGTTCATGGCTACTGGTTCCGGGAAGGGGCCAATACAGACCAGGATGCTCC  
AGTGGCCACAAACAACCCAGCTCGGGCAGTGTGGGAGGAGACTCGGGACCGATTCCACCTCCTTG  
GGGACCCACATACCAAGAATTGCACCCTGAGCATCAGAGATGCCAGAAGAAGTGATGCGGGGAGA  
TACTTCTTTCGTATGGAGAAAGGAAGTATAAAATGGAATTATAAACATCACCGGCTCTCTGTGAA  
TGTGACAGCCTTGACCCACAGGCCCAACATCCTCATCCCAGGCACCCTGGAGTCCGGCTGCCCCC  
AGAATCTGACCTGCTCTGTGCCCTGGGCCTGTGAGCAGGGGACACCCCCTATGATCTCCTGGATA  
GGGACCTCCGTGTCCCCCTGGACCCCTCCACCACCCGCTCCTCGGTGCTCACCTCATCCCACA  
GCCCCAGGACCATGGCACCAGCCTCACCTGTCAGGTGACCTTCCCTGGGGCCAGCGTGACCACGA  
ACAAGACCGTCCATCTCAACGTGTCTACCCGCCTCAGAACTTGACCATGACTGTCTTCCAAGGA  
GACGGCACAGTATCCACAGTCTTGGGAAATGGCTCATCTCTGTCACTCCCAGAGGGCCAGTCTCT  
GCGCCTGGTCTGTGCAGTTGATGCAGTTGACAGCAATCCCCCTGCCAGGCTGAGCCTGAGCTGGA  
GAGGCCTGACCCTGTGCCCTCACAGCCCTCAAACCCGGGGGTGCTGGAGCTGCCTTGGGTGCAC  
CTGAGGGATGCAGCTGAATTCACCTGCAGAGCTCAGAACCTCTCGGCTCTCAGCAGGTCTACCT  
GAACGTCTCCCTGCAGAGCAAAGCCACATCAGGAGTGACTCAGGGGGTGGTCGGGGGAGCTGGAG  
CCACAGCCCTGGTCTTCCTGTCCTTCTGCGTCATCTTCGTTGTAGTGAGGTCTGCAGGAAGAAA  
TCGGCAAGGCCAGCAGCGGGCGTGGGAGATACGGGCATAGAGGATGCAAACGCTGTCAGGGGTTC  
AGCCTCTCAGGGGCCCCTGACTGAACCTTGGGCAGAAGACAGTCCCCCAGACCAGCCTCCCCCAG  
CTTCTGCCCCTCCTCAGTGGGGGAAGGAGAGCTCCAGTATGCATCCCTCAGCTTCCAGATGGTG  
AAGCCTTGGGACTCGCGGGGACAGGAGGCCACTGACACCGAGTACTCGGAGATCAAGATCCACAG  
ATGAGAACTGCAGAGACTCACCTGATTGAGGGATCACAGCCCCTCCAGGCAAGGGAGAAGTCA  
GAGGCTGATTCTTGTAGAATTAACAGCCCTCAACGTGATGAGCTATGATAACACTATGAATTATG  
TGCAGAGTGAAAAGCACACAGGCTTTAGAGTCAAAGTATCTCAAACCTGAATCCACACTGTGCCC  
TCCCTTTTATTTTTTTAACTAAAAGACAGACAAATTCCTA

## **FIGURE 86**

MLLLLLPLLWGRERAEGQTSKLLTMQSSVTVQEGLCVHVPCSFSYPSHGWIYPGPVVHGYWFREG  
ANTDQDAPVATNNPARAVWEETRDRFHLLGDPHTKNCTLSIRDARRSDAGRYFFRMEKGSIKWNY  
KHHRLSVNVTALTHRPNILIPGTLESGCPQNLTCVWPWACEQGTPPMISWIGTSVSPLDPSTTRS  
SVLTLPQPQDHGTSLTCQVTFPGASVTTNKTVHLNVSYPPQNLTMTVFQGDGTVSTVLGNGSSL  
SLPEGQSLRLVCAVDAVDSNPPARLSLSWRGLTLCPSQPSNPGVLELPWVHLRDAAEFTCRAQNP  
LGSQQVYLVNLSLQSKATSGVTQGVVGGAGATALVFLSFCVIFVVVRSCRKKSARPAAGVGDGTGIE  
DANAVRGSSASQGPLEPWAEDSPPDQPPPASARSSVGEGELQYASLSFQMVKPWDSRGQEATDTE  
YSEIKIHR

**Signal peptide:**

amino acids 1-15

**Transmembrane domain:**

amino acids 351-370

**FIGURE 87**

AGAAAGCTGCACTCTGTTGAGCTCCAGGGCGCAGTGGAGGGAGGGAGTGAAGGAGCTCTCTGTAC  
CCAAGGAAAGTGCAGCTGAGACTCAGACAAGATTACAATGAACCAACTCAGCTTCCTGCTGTTTC  
TCATAGCGACCACCAGAGGATGGAGTACAGATGAGGCTAATACTTACTTCAAGGAATGGACCTGT  
TCTTCGTCTCCATCTCTGCCCAGAAGCTGCAAGGAAATCAAAGACGAATGTCCTAGTGCATTTGA  
TGGCCTGTATTTTCTCCGCACTGAGAATGGTGTATCTACCAGACCTTCTGTGACATGACCTCTG  
GGGTGGCGGCTGGACCCTGGTGGCCAGCGTGCATGAGAATGACATGCGTGGGAAGTGCACGGTG  
GGCGATCGCTGGTCCAGTCAGCAGGGCAGCAAAGCAGACTACCCAGAGGGGGACGGCAACTGGGC  
CAACTACAACACCTTTGGATCTGCAGAGGCGGCCACGAGCGATGACTACAAGAACCCTGGCTACT  
ACGACATCCAGGCCAAGGACCTGGGCATCTGGCACGTGCCCAATAAGTCCCCCATGCAGCACTGG  
AGAAACAGCTCCCTGCTGAGGTACCGCACGGACACTGGCTTCCTCCAGACACTGGGACATAATCT  
GTTTGGCATCTACCAGAAATATCCAGTGAAATATGGAGAAGGAAAGTGTTGGACTGACAACGGCC  
CGGTGATCCCTGTGGTCTATGATTTTGGCGACGCCCAGAAAACAGCATCTTATTACTCACCCCTAT  
GGCCAGCGGGAATTCAGTGCGGGATTTGTTTCAGTTCAGGGTATTTAATAACGAGAGAGCAGCCAA  
CGCCTTGTGTGCTGGAATGAGGGTCACCGGATGTAACTGAGCATCACTGCATTGGTGGAGGAG  
GATACTTTCCAGAGGCCAGTCCCCAGCAGTGTGGAGATTTTCTGGTTTTGATTGGAGTGGATAT  
GGAACATGTTGGTTACAGCAGCAGCCGTGAGATAACTGAGGCAGCTGTGCTTCTATTCTATCG  
TTGAGAGTTTTGTGGGAGGGAACCCAGACCTCTCCTCCCAACCATGAGATCCCAAGGATGGAGAA  
CAACTTACCCAGTAGCTAGAATGTTAATGGCAGAAGAGAAAACAATAAATCATATTGACTCAAGA  
AAAAAA

## **FIGURE 88**

MNQLSFLLFLIATTRGWSTDEANTYFKEWTCSSSPSLPRSCKEIKDECPSAFDGLYFLRTENGVI  
YQTFCDMTSGGGGWTLVASVHENDMRGKCTVGDRWSSQQGSKADYPEGDGNWANYNTFGSAEAAT  
SDDYKNPGYYDIQAKDLGIWHVPNKSPMQHWRNSSLLRYRTDTGFLQTLGHNLFGIYQKYPVKYG  
EGKCWTDNGPVI PVVYDFGDAQKTASYSPYGQREFTAGFVQFRVFNNERAANALCAGMRVTGCN  
TEHHCIGGGGYFPEAS PQQCGDFSGFDWSGYGTHVGYSSSREITEAAVLLFYR

**Important features:**

**Signal peptide:**

amino acids 1-16

**N-glycosylation site.**

amino acids 163-167

**Glycosaminoglycan attachment sites.**

amino acids 74-78, 289-293

**N-myristoylation sites.**

amino acids 76-82, 115-121, 124-130, 253-259, 292-298

**FIGURE 89**

CTAGATTTGTCGGCTTGCGGGGAGACTTCAGGAGTCGCTGTCTCTGAACTTCCAGCCTCAGAGAC  
CGCCGCCCTTGTCCTCCGAGGGCC**ATG**GGCCGGGTCTCAGGGCTTGCGCCCTCTCGCTTCCTGACG  
CTCCTGGCGCATCTGGTGGTCGTCATCACCTTATTCTGGTCCCGGGACAGCAACATACAGGCCTG  
CCTGCCCTCTCACGTTACCCCCGAGGAGTATGACAAGCAGGACATTCAGCTGGTGGCCGCGCTCT  
CTGTCACCCTGGGCCTCTTTGCAGTGGAGCTGGCCGGTTTCCTCTCAGGAGTCTCCATGTTCAAC  
AGCACCCAGAGCCTCATCTCCATTGGGGCTCACTGTAGTGCATCCGTGGCCCTGTCCTTCTTCAT  
ATTCGAGCGTTGGGAGTGCACTACGTATTGGTACATTTTTGTCTTCTGCAGTGCCCTTCCAGCTG  
TCACTGAAATGGCTTTATTCGTCACCGTCTTTGGGCTGAAAAAGAAACCCTTCT**TGA**TTACCTTCA  
TGACGGGAACCTAAGGACGAAGCCTACAGGGGCAAGGGCCGCTTCGTATTCCTGGAAGAAGGAAG  
GCATAGGCTTCGGTTTTCCCTCGGAAACTGCTTCTGCTGGAGGATATGTGTTGGAATAATTACG  
TCTTGAGTCTGGGATTATCCGCATTGTATTTAGTGCTTTGTAATAAAATATGTTTTGTAGTAACA  
TTAAGACTTATATACAGTTTTAGGGGACAATTAAAAAAAAAAAA

## **FIGURE 90**

MGRVSGLVPSRFLTLLAHLVVVITLFWSRDSNIQACLPLTFTPEEYDKQDIQLVAALSVTLGLFA  
VELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFCSALPAVTEMALFV  
TVFGLKKKPF

**Transmembrane domain:**

amino acids 12-28 (type II), 51-66, 107-124

**FIGURE 91**

CTGGGACCCCGAAAAGAGAAGGGGAGAGCGAGGGGACGAGAGCGGAGGAGGAAGATGCAACTGAC  
TCGCTGCTGCTTCGTGTTCTTGGTGCAGGGTAGCCTCTATCTGGTCATCTGTGGCCAGGATGATG  
GTCCTCCCGGCTCAGAGGACCCTGAGCGTGATGACCACGAGGGCCAGCCCCGGCCCCGGGTGCCT  
CGGAAGCGGGGCCACATCTCACCTAAGTCCCGCCCCATGGCCAATTCCACTCTCCTAGGGCTGCT  
GGCCCCGCCTGGGGAGGCTTGGGGCATTCTTGGGCAGCCCCCAACCGCCCCGAACCACAGCCCCC  
CACCTCAGCCAAGGTGAAGAAAATCTTTGGCTGGGGCGACTTCTACTCCAACATCAAGACGGTG  
GCCCTGAACCTGCTCGTCACAGGGAAGATTGTGGACCATGGCAATGGGACCTTCAGCGTCCACTT  
CCAACACAATGCCACAGGCCAGGGAACATCTCCATCAGCCTCGTGCCCCCAGTAAAGCTGTAG  
AGTTCCACCAGGAACAGCAGATCTTCATCGAAGCCAAGGCCTCCAAAATCTTCAACTGCCGGATG  
GAGTGGGAGAAGGTAGAACGGGGCCGCCGACCTCGCTTTGCACCCACGACCCAGCCAAGATCTG  
CTCCCGAGACCACGCTCAGAGCTCAGCCACCTGGAGCTGCTCCAGCCCTTCAAAGTCGTCTGTG  
TCTACATCGCCTTCTACAGCACGGACTATCGGCTGGTCCAGAAGGTGTGCCCAGATTACAACCTAC  
CATAGTGATACCCCCTACTACCCATCTGGGTGACCCGGGGCAGGCCACAGAGGCCAGGCCAGGGC  
TGGAAGGACAGGCCTGCCCATGCAGGAGACCATCTGGACACCGGGCAGGGAAGGGGTGGGCCTC  
AGGCAGGGAGGGGGGTGGAGACGAGGAGATGCCAAGTGGGGCCAGGGCCAAGTCTCAAGTGGCAG  
AGAAAGGGTCCCAAGTGCTGGTCCCAACCTGAAGCTGTGGAGTGACTAGATCACAGGAGCACTGG  
AGGAGGAGTGGGCTCTCTGTGCAGCCTCACAGGGCTTTGCCACGGAGCCACAGAGAGATGCTGGG  
TCCCCGAGGCCTGTGGGCAGGCCGATCAGTGTGGCCCCAGATCAAGTCATGGGAGGAAGCTAAGC  
CCTTGGTTCTTGCCATCCTGAGGAAAGATAGCAACAGGGAGGGGGAGATTTTCATCAGTGTGGACA  
GCCTGTCAACTTAGGATGGATGGCTGAGAGGGCTTCCCTAGGAGCCAGTCAGCAGGGTGGGGTGGG  
GCCAGAGGAGCTCTCCAGCCCTGCCTAGTGGGCGCCCTGAGCCCCTTGTCGTGTGCTGAGCATGG  
CATGAGGCTGAAGTGGCAACCCTGGGGTCTTTGATGTCTTGACAGATTGACCATCTGTCTCCAGC  
CAGGCCACCCCTTTCCAAAATTCCTCTTCTGCCAGTACTCCCCCTGTACCACCCATTGCTGATG  
GCACACCCATCCTTAAGCTAAGACAGGACGATTGTGGTCCTCCACACTAAGGCCACAGCCCATC  
CGCGTGCTGTGTGTCCCTCTTCCACCCCCAACCCTGCTGGCTCCTCTGGGAGCATCCATGTCCCG  
GAGAGGGGTCCCTCAACAGTCAGCCTCACCTGTCAGACCGGGGTCTCCCGGATCTGGATGGCGC  
CGCCCTCTCAGCAGCGGGCACGGGTGGGGCGGGGCCGGGCCGAGAGCATGTGCTGGATCTGTTC  
TGTGTGTCTGTCTGTGGGTGGGGGAGGGGAGGGAAGTCTTGTGAAACCGCTGATTGCTGACTTT  
TGTGTGAAGAATCGTGTTCTTGGAGCAGGAAATAAAGCTTGCCCCGGGGCA

92/168

## **FIGURE 92**

MQLTRCCFVFLVQGSLYLVICGQDDGPPGSEDPERDDHEGQPRPRVPRKRGHISPKSRPMANSTL  
LGLLAPPGEAWGILGQPPNRPNHSPPPSAKVKKIFGWGDFYSNIKTVALNLLVTGKIVDHGNGTF  
SVHFQHNATGQGNISISLVPPSKAVEFHQEQQIFIEAKASKIFNCRMWEKVERGRRTSLCTHDP  
AKICSRDHAQSSATWSCSQPFKVVCVYIAFYSTDYRLVQKVCPDYNHSDTPYYPSG

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-14

#### **N-glycosylation sites.**

amino acids 62-65, 127-130, 137-140, 143-146

#### **2-oxo acid dehydrogenases acyltransferase**

amino acids 61-71



**FIGURE 93**

CGGTGGCCATGACTGCGGCCGTGTTCTTCGGCTGCGCCTTCATTGCCTTCGGGCCTGCGCTCGCC  
CTTTATGTCTTCACCATCGCCATCGAGCCGTTGCGTATCATCTTCCTCATCGCCGGAGCTTTCTT  
CTGGTTGGTGTCTCTACTGATTTTCGTCCCTTGTTTGGTTCATGGCAAGAGTCATTATTGACAACA  
AAGATGGACCAACACAGAAATATCTGCTGATCTTTGGAGCGTTTGTCTCTGTCTATATCCAAGAA  
ATGTTCCGATTTGCATATTATAAACTCTTAAAAAAGCCAGTGAAGGTTTGAAGAGTATAAACCC  
AGGTGAGACAGCACCTCTATGCGACTGCTGGCCTATGTTTCTGGCTTGGGCTTTGGAATCATGA  
GTGGAGTATTTTCCTTTGTGAATACCCTATCTGACTCCTTGGGGCCAGGCACAGTGGGCATTCAT  
GGAGATTCTCCTCAATTCTTCCTTTATTCAGCTTTCATGACGCTGGTCATTATCTTGCTGCATGT  
ATTCTGGGGCATTGTATTTTTTGGATGGCTGTGAGAAGAAAAAGTGGGGCATCCTCCTTATCGTTC  
TCCTGACCCACCTGCTGGTGTGAGCCAGACCTTCATAAGTTCTTATTATGGAATAAACCTGGCG  
TCAGCATTTATAATCCTGGTGCTCATGGGCACCTGGGCATTCTTAGCTGCGGGAGGCAGCTGCCG  
AAGCCTGAAACTCTGCCTGCTCTGCCAAGACAAGAACTTTCTTCTTTACAACCAGCGCTCCAGATT  
AACCTCAGGGAACCAGCACTTCCCAAACCGCAGACTACATCTTTAGAGGAAGCACAACTGTGCCT  
TTTTCTGAAAATCCCTTTTTCTGGTGGAATTGAGAAAGAAATAAACTATGCAGATA

## **FIGURE 94**

MTAAVFFGCAFI AFGPALALYVFTIAIEPLRIIFLIAGAFFWLVSLLISSLVWF MARVIIDNKDG  
PTQKYLLIFGAFVSVYIQEMFRFAYYKLLKKASEGLKSINPGETAPSMRLLAYVSGLGFGIMSGV  
ESFVNTLSDSLGP GTVGIHG DSPQFFLYSAFMTLVIILLHVFWGIVFFDGCEKKKWGILLIVLLT  
HLLVSAQTFISSYYGINLASAFIILVLMGTWAF LAAGGSCRSLKLCLLCQDKNFLLYNQRSR

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-19

#### **Transmembrane domains:**

amino acids 32-51, 119-138, 152-169, 216-235

#### **Glycosaminoglycan attachment site.**

amino acids 120-123

#### **Sodium:neurotransmitter symporter family protein**

amino acids 31-65

**FIGURE 95**

AATTTTTCACCAGAGTAAACTTGAGAAACCAACTGGACCTTGAGTATTGTACATTTTGCCTCGTG  
GACCCAAAGGTAGCAATCTGAAAC**ATG**AGGAGTACGATTCTACTGTTTTGTCTTCTAGGATCAAC  
TCGGTCATTACCACAGCTCAAACCTGCTTTGGGACTCCCTCCCACAAAACCTGGCTCCGGATCAGG  
GAACACTACCAAACCAACAGCAGTCAAATCAGGTCTTTCCTTCTTTAAGTCTGATACCATTAACA  
CAGATGCTCACACTGGGGCCAGATCTGCATCTGTAAATCCTGCTGCAGGAATGACACCTGGTAC  
CCAGACCCACCCATTGACCCTGGGAGGGTTGAATGTACAACAGCAACTGCACCCACATGTGTTAC  
CAATTTTTGTACACAACCTTGGAGCCCAGGGCACTATCCTAAGCTCAGAGGAATTGCCACAAATC  
TTCACGAGCCTCATCATCCATTCCCTTGTTCCCGGGAGGCATCCTGCCCACCAGTCAGGCAGGGGC  
TAATCCAGATGTCCAGGATGGAAGCCTTCCAGCAGGAGGAGCAGGTGTAAATCCTGCCACCCAGG  
GAACCCAGCAGGCCGCCCTCCCAACTCCCAGTGGCACAGATGACGACTTTGCAGTGACCACCCCT  
GCAGGCATCCAAAGGAGCACACATGCCATCGAGGAAGCCACCACAGAATCAGCAAATGGAATTCA  
**GTAAG**CTGTTTCAAATTTTTTCAACTAAGCTGCCTCGAATTTGGTGATACATGTGAATCTTTATC  
ATTGATTATATTATGGAATAGATTGAGACACATTGGATAGTCTTAGAAGAAATTAATTCTTAATT  
TACCTGAAAATATTCTTGAAATTCAGAAAATATGTTCTATGTAGAGAATCCCAACTTTTAAAAA  
CAATAATTCAATGGATAAATCTGTCTTTGAAATATAACATTATGCTGCCTGGATGATATGCATAT  
TAAAACATATTTGGAAAACCTGGAAA  
AAA

## **FIGURE 96**

MRSTILLFCLLGSTRSLPQLKPALGLPPTKLAPDQGTLPNQQQSNQVFPSLSLIPLTQM  
LTLGPDHLHLLNPAAGMTPGTQTHPLTLGGLNVQQQLHPHVLPIFVTQLGAQGTTLSSEE  
LPQIFTSLIIHSLFPGGILPTSQAGANPDVQDGSLPAGGAGVNPATQGTPAGRLPTPSG  
TDDDFAVTTPAGIQRSTHAIEEATTESANGIQ

**Signal peptide:**

amino acids 1-16

**FIGURE 97**

GCTCAAGTGCCCTGCCTTGCCCCACCCAGCCCAGCCTGGCCAGAGCCCCCTGGAGAAGGAGCTCT  
 CTTCTTGCTTGGCAGCTGGACCAAGGGAGCCAGTCTTGGGCGCTGGAGGGCCTGTCCTGACC**ATG**  
 GTCCCTGCCTGGCTGTGGCTGCTTTGTGTCTCCGTCCCCCAGGCTCTCCCCAAGGCCAGCCTGC  
 AGAGCTGTCTGTGGAAGTTCCAGAAAACATATGGTGGAAATTTCCCTTTATACCTGACCAAGTTGC  
 CGCTGCCCCGTGAGGGGGCTGAAGGCCAGATCGTGCTGTCAGGGGACTCAGGCAAGGCAACTGAG  
 GGCCCATTTGCTATGGATCCAGATTCTGGCTTCCTGCTGGTGACCAGGGCCCTGGACCGAGAGGA  
 GCAGGCAGAGTACCAGCTACAGGTCACCCTGGAGATGCAGGATGGACATGTCTTGTGGGGTCCAC  
 AGCCTGTGCTTGTGCACGTGAAGGATGAGAATGACCAGGTGCCCCATTTCTCTCAAGCCATCTAC  
 AGAGCTCGGCTGAGCCGGGGTACCAGGCCTGGCATCCCCCTTCCTCTTCCTTGAGGCTTCAGACCG  
 GGATGAGCCAGGCACAGCCAACTCGGATCTTCGATTCCACATCCTGAGCCAGGCTCCAGCCCAGC  
 CTTCCCCAGACATGTTCCAGCTGGAGCCTCGGCTGGGGGCTCTGGCCCTCAGCCCCAAGGGGAGC  
 ACCAGCCTTGACCACGCCCTGGAGAGGACCTACCAGCTGTTGGTACAGGTCAAGGACATGGGTGA  
 CCAGGCCTCAGGCCACCAGGCCACTGCCACCGTGGAAGTCTCCATCATAGAGAGCACCTGGGTGT  
 CCCTAGAGCCTATCCACCTGGCAGAGAATCTCAAAGTCTTATACCCGCACCACATGGCCCAGGTA  
 CACTGGAGTGGGGGTGATGTGCACTATCACCTGGAGAGCCATCCCCCGGGACCCTTTGAAGTGAA  
 TGCAGAGGGAAACCTCTACGTGACCAGAGAGCTGGACAGAGAAGCCCAGGCTGAGTACCTGCTCC  
 AGGTGCGGGCTCAGAATTCCCATGGCGAGGACTATGCGGCCCTCTGGAGCTGCACGTGCTGGTG  
 ATGGATGAGAATGACAACGTGCCTATCTGCCCTCCCCGTGACCCACAGTCAGCATCCCTGAGCT  
 CAGTCCACCAGGTACTGAAGTGACTAGACTGTCAGCAGAGGATGCAGATGCCCCCGGCTCCCCCA  
 ATTCCCACGTTGTGTATCAGCTCCTGAGCCCTGAGCCTGAGGATGGGGTAGAGGGGAGAGCCTTC  
 CAGGTGGACCCCACTTCAGGCAGTGTGACGCTGGGGGTGCTCCCACTCCGAGCAGGCCAGAACAT  
 CCTGCTTCTGGTGCTGGCCATGGACCTGGCAGGCGCAGAGGGTGGCTTCAGCAGCACGTGTGAAG  
 TCGAAGTCGCAGTCACAGATATCAATGATCACGCCCTGAGTTCATCACTTCCCAGATTGGGCCT  
 ATAAGCCTCCCTGAGGATGTGGAGCCCGGACTCTGGTGGCCATGCTAACAGCCATTGATGCTGA  
 CCTCGAGCCCGCCTTCCGCCTCATGGATTTTGCCATTGAGAGGGGAGACACAGAAGGGACTTTTG  
 GCCTGGATTGGGAGCCAGACTCTGGGCATGTTAGACTCAGACTCTGCAAGAACCTCAGTTATGAG  
 GCAGCTCCAAGTCATGAGGTGGTGGTGGTGGTGCAGAGTGTGGCGAAGCTGGTGGGGCCAGGCCC  
 AGGCCCTGGAGCCACCGCCACGGTGACTGTGCTAGTGGAGAGAGTGATGCCACCCCCCAAGTTGG  
 ACCAGGAGAGCTACGAGGCCAGTGTCCCCATCAGTGCCCCAGCCGGCTCTTTCTGCTGACCATC  
 CAGCCCTCCGACCCCATCAGCCGAACCCTCAGGTTCTCCCTAGTCAATGACTCAGAGGGCTGGCT  
 CTGCATTGAGAAATTCTCCGGGGAGGTGCACACCGCCCAGTCCCTGCAGGGCGCCCAGCCTGGGG  
 ACACCTACACGGTGCTTGTGGAGGCCAGGATACAGCCCTGACTCTTGCCCCTGTGCCCTCCCCA  
 TACCTCTGCACACCCCGCCAAGACCATGGCTTGATCGTGAGTGGACCCAGCAAGGACCCCGATCT  
 GGCCAGTGGGCACGGTCCCTACAGCTTCACCCTTGGTCCCAACCCACGGTGCAACGGGATTGGC  
 GCCTCCAGACTCTCAATGGTTCCCATGCCTACCTCACCTTGGCCCTGCATTGGGTGGAGCCACGT  
 GAACACATAATCCCCGTGGTGGTCAGCCACAATGCCAGATGTGGCAGCTCCTGGTTTCGAGTGAT  
 CGTGTGTCGCTGCAACGTGGAGGGGCAGTGCATGCGCAAGGTGGGCGCATGAAGGGCATGCCCA  
 CGAAGCTGTCGGCAGTGGGCATCCTTGTTAGGCACCCTGGTAGCAATAGGAATCTTCCTCATCCTC  
 ATTTTACCCACTGGACCATGTCAAGGAAGAAGGACCCGGATCAACCAGCAGACAGCGTGCCCCCT  
 GAAGGCGACTGT**TGA**ATGGCCCAGGCAGCTCTAGCTGGGAGCTTGCCCTCTGGCTCCATCTGAG  
 TCCCCTGGGAGAGAGCCCAGCACCCAAGATCCAGCAGGGGACAGGACAGAGTAGAAGCCCCCTCCA  
 TCTGCCCTGGGGTGGAGGCACCATCACCATCACAGGCATGTCTGCAGAGCCTGGACACCAACTT  
 TATGGACTGCCCATGGGAGTGCTCCAAATGTCAGGGTGTTTGCCCAATAATAAAGCCCCAGAGAA  
 CTGGGCTGGGCCCTATGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAG

**FIGURE 98**

MVPAWLWLLCVSVPQALPKAQPAELSVVPEVNYGGNFPLYLTKLPLPREGAEGQIVLSGDSGKAT  
EGPFAMDPDSGFLLVTRALDREEQAEYQLQVTLEMQDGHVLWGPQPVLVHVKDENDQVPHFSQAI  
YRARLSRGTRPGIPFLFLEASDRDEPGTANSDLRFHILSQAPAQSPDMFQLEPRLGALALSPKG  
STSLDHALERTYQLLVQVKDMGDQASGHQATATVEVSIIESTWVSLEPIHLAENLKVLYPHHMAQ  
VHWSGGDVHYHLESHPPGPFVNAEGNLYVTRELDREAQAEYLLQVRAQNSHGEDIAAPLELHVL  
VMDENDNVPICPPRDPTVSIPELSPPGTEVTRLSAEDADAPGSPNSHVYQLLSPEPEDGVEGRA  
FQVDPTSGSVTLGVLPLRAGQNILLVLAMDLAGAEGGFSSTCEVEVAVTDINDHAPEFITSQIG  
PISLPEDVEPGTLVAMLTADLEPAFRLMDFAIERGDTEGTFGLDWEPDSGHVRLRLCKNLSY  
EAAPSHEVVVVVQSVAKLVGPGPGGATATVTVLVERVMPPPKLDQESYEASVPISAPAGSFLLT  
IQPSDPISRTLRFSLVNDSEGWLCIEKFSGEVHTAQSLQGAQPGDTYTVLVEAQDTALTAPVPS  
QYLCTPRQDHGLIVSGPSKDPDLASGHGPYSFTLGPNPTVQRDWRLQTLNGSHAYLTALHWVEP  
REHIIPVVVSHNAQMWQLLVRVIVCRCNVEGQCMRKVGRMKGMPTKLSAVGILVGTTLVAIGIFLI  
LIFTHWTMSRKKDPDQPADSVPLKATV

**Signal peptide:**

amino acids 1-18

**Transmembrane domain:**

amino acids 762-784

**FIGURE 99**

GGCTGACCGTGCTACATTGCCTGGAGGAAGCCTAAGGAACCCAGGCATCCAGCTGCCCACGCCTG  
AGTCCAAGATTCTTCCCAGGAACACAAACGTAGGAGACCCACGCTCCTGGAAGCACCAGCCTTTA  
TCTCTTCACCTTCAAGTCCCCTTTCTCAAGAATCCTCTGTTCTTTGCCCTCTAAAGTCTTGGTAC  
ATCTAGGACCCAGGCATCTTGCTTTCCAGCCACAAAGAGACAGATGAAAGATGCAGAAAGGAAATG  
TTCTCCTTATGTTTGGTCTACTATTGCATTTAGAAGCTGCAACAAATTCCAATGAGACTAGCACC  
TCTGCCAACACTGGATCCAGTGTGATCTCCAGTGGAGCCAGCACAGCCACCAACTCTGGGTCCAG  
TGTGACCTCCAGTGGGGTCAGCACAGCCACCATCTCAGGGTCCAGCGTGACCTCCAATGGGGTCA  
GCATAGTCACCAACTCTGAGTTCATACAACCTCCAGTGGGATCAGCACAGCCACCAACTCTGAG  
TTCAGCACAGCGTCCAGTGGGATCAGCATAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGG  
GGCCAGCACAGCCACCAACTCTGAGTCCAGCACACCCTCCAGTGGGGCCAGCACAGTCACCAACT  
CTGGGTCCAGTGTGACCTCCAGTGGAGCCAGCACTGCCACCAACTCTGAGTCCAGCACAGTGTCC  
AGTAGGGCCAGCACTGCCACCAACTCTGAGTCTAGCACACTCTCCAGTGGGGCCAGCACAGCCAC  
CAACTCTGACTCCAGCACAACTCCAGTGGGGCTAGCACAGCCACCAACTCTGAGTCCAGCACAA  
CCTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAGTGTCCAGTAGGGCCAGCACT  
GCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAG  
AACGACCTCCAATGGGGCTGGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGGGGCCA  
GCACAGCCACCAACTCTGACTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAG  
TCCAGCACGACCTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGG  
GGCTAGCACAGCCACCAACTCTGACTCCAGCACAACTCCAGTGGGGCCGGCACAGCCACCAACT  
CTGAGTCCAGCACAGTGTCCAGTGGGATCAGCACAGTCACCAATTCTGAGTCCAGCACACCCTCC  
AGTGGGGCCAACACAGCCACCAACTCTGAGTCCAGTACGACCTCCAGTGGGGCCAACACAGCCAC  
CAACTCTGAGTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAGTCCAGCACAA  
CCTCCAGTGGGGTCAGCACAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCTAGCACA  
GCCACCAACTCTGACTCCAGCACAACTCCAGTGAGGCCAGCACAGCCACCAACTCTGAGTCTAG  
CACAGTGTCCAGTGGGATCAGCACAGTCACCAATTCTGAGTCCAGCACAACTCCAGTGGGGCCA  
ACACAGCCACCAACTCTGGGTCCAGTGTGACCTCTGCAGGCTCTGGAACAGCAGCTCTGACTGGA  
ATGCACACAACTTCCCATAGTGCATCTACTGCAGTGAAGTGAAGCAAGCCTGGTGGGTCCCTGGT  
GCCGTGGGAAATCTTCCTCATCACCTGGTCTCGGTTGTGGCGGCCGTGGGGCTCTTTGCTGGGC  
TCTTCTTCTGTGTGAGAAACAGCCTGTCCCTGAGAAACACCTTTAACACAGCTGTCTACCACCCT  
CATGGCCTCAACCATGGCCTTGGTCCAGGCCCTGGAGGGAATCATGGAGCCCCCACAGGCCAG  
GTGGAGTCTAAGTGGTTCTGGAGGAGACCAGTATCATCGATAGCCATGGAGATGAGCGGGAGGA  
ACAGCGGGCCCTTGAGCAGCCCCGGAAGCAAGTGCCGCATTCTTCAGGAAGGAAGAGACCTGGGCA  
CCCAAGACCTGGTTTTCCTTTTCATTCATCCCAGGAGACCCCTCCAGCTTTGTTTGAGATCCTGAA  
AATCTTGAAGAAGGTATTCCTCACCTTTCTTGCCTTTACCAGACACTGGAAGAGAATACTATAT  
TGCTCATTTAGCTAAGAAATAAATACATCTCATCTAACACACACGACAAAGAGAAGCTGTGCTTG  
CCCCGGGGTGGGTATCTAGCTCTGAGATGAACTCAGTTATAGGAGAAAACCTCCATGCTGGACTC  
CATCTGGCATTCAAATCTCCACAGTAAATCCAAAGACCTCAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 100**

MKMQKGNVLLMFGLLLHLEAATNSNETSTSANTGSSVISSGASTATNSGSSVTSSGVSTATISGS  
SVTSNGVSIVTNSEFHTTSSGISTATNSEFSTASSGISIATNSESSTTSSGASTATNSESSTPSS  
GASTVTNSGSSVTSSGASTATNSESSTVSSRASTATNSESSTLSSGASTATNSDSSTTSSGASTA  
TNSESSTTSSGASTATNSESSTVSSRASTATNSESSTTSSGASTATNSESRTTSNGAGTATNSES  
STTSSGASTATNSDSSTVSSGASTATNSESSTTSSGASTATNSESSTTSSGASTATNSDSSTTSS  
GAGTATNSESSTVSSGISTVTNSESSTPSSGANTATNSESSTTSSGANTATNSESSTVSSGASTA  
TNSESSTTSSGVSTATNSESSTTSSGASTATNSDSSTTSSEASTATNSESSTVSSGISTVTNSES  
STTSSGANTATNSGSSVTSAGSGTAALTGMHTTSHSASTAVSEAKPGGSLVPWEIFLITLVSVVA  
AVGLFAGLFFCVRNSLSLRNTFNTAVYHPHGLNHGLGPGPGGNHGAPHRPRWSPNWFWRPVSII  
AMEMSGRNSGP

**Signal peptide:**

amino acids 1-20

**Transmembrane domain:**

amino acids 510-532



**FIGURE 101**

GGCCGGACGCCTCCGCGTTACGGGATGAATTAACGGCGGGTTCCGCACGGAGGTTGTGACCCCTA  
CGGAGCCCCAGCTTGCCACGCACCCCACTCGGCGTCGCGCGGCGTGCCCTGCTTGTCACAGGTG  
GGAGGCTGGAACATATCAGGCTGAAAAACAGAGTGGGTACTCTCTTCTGGGAAGCTGGCAACAAAT  
GGATGATGTGATAT**ATG**CATTCCAGGGGAAGGGAAATTGTGGTGCTTCTGAACCCATGGTCAATT  
AACGAGGCAGTTTCTAGCTACTGCACGTACTTCATAAAGCAGGACTCTAAAAGCTTTGGAATCAT  
GGTGTCATGGAAAGGGATTTACTTTATACTGACTCTGTTTTGGGGAAGCTTTTTTGGGAAGCATTT  
TCATGCTGAGTCCCTTTTTACCTTTGATGTTTGTAACCCATCTTGGTATCGCTGGATCAACAAC  
CGCCTTGTTGGCAACATGGCTCACCTACCTGTGGCATTATTGGAGACCATGTTTGGTGTAAGT  
GATTATAACTGGGGATGCATTTGTTCTGGAGAAAGAAGTGTCATTATCATGAACCATCGGACAA  
GAATGGACTGGATGTTCTGTGGAATTGCCTGATGCGATATAGCTACCTCAGATTGGAGAAAATT  
TGCCTCAAAGCGAGTCTCAAAGGTGTTCTGGATTTGGTTGGGCCATGCAGGCTGCTGCCTATAT  
CTTCATTCATAGGAAATGGAAGGATGACAAGAGCCATTTCTGAAGACATGATTGATTACTTTTGTG  
ATATTCACGAACCACTTCAACTCCTCATATTCAGAGAGGGACTGATCTCACAGAAAACAGCAAG  
TCTCGAAGTAATGCATTTGCTGAAAAAATGGACTTCAGAAATATGAATATGTTTTACATCCAAG  
AACTACAGGCTTTACTTTTGTGGTAGACCGTCTAAGAGAAGGTAAGAACCTTGATGCTGTCCATG  
ATATCACTGTGGCGTATCCTCACAACATTCCTCAATCAGAGAAGCACCTCCTCCAAGGAGACTTT  
CCCAGGGAAATCCACTTTACGTCCACCGGTATCCAATAGACACCCTCCCCACATCCAAGGAGGA  
CCTTCAACTCTGGTGCCACAAACGGTGGGAAGAGAAAGAAGAGAGGCTGCGTTCCTTCTATCAAG  
GGGAGAAGAATTTTTATTTTACCGGACAGAGTGTCATTCCACCTTGCAAGTCTGAACTCAGGGTC  
CTTGTTGGTCAAATTGCTCTCTATACTGTATTGGACCCTGTTACGCCCTGCAATGTGCCTACTCAT  
ATATTTGTACAGTCTTGTTAAGTGGTATTTTATAATCACCATTGTAATCTTTGTGCTGCAAGAGA  
GAATATTTGGTGGACTGGAGATCATAGAATTGCATGTTACCGACTTTTACACAAACAGCCACAT  
TTAAATTCAAAGAAAAATGAG**TAA**GATTATAAGGTTTGCCATGTGAAAACCTAGAGCATATTTTG  
GAAATGTTCTAAACCTTTCTAAGCTCAGATGCATTTTTGCATGACTATGTGCAATATTTCTTACT  
GCCATCATTATTTGTTAAAGATATTTTGCCTTAATTTTGTGGGAAAAATATTGCTACAATTTTT  
TTAATCTCTGAATGTAATTTGATACTGTGTACATAGCAGGGAGTGATCGGGGTGAAATAACTT  
GGCCAGAATATTATTAAACAATCATCAGGCTTTTAAA

## **FIGURE 102**

MHSRGREIVVLLNPWSINEAVSSYCTYFIKQDSKSFSGIMVSWKGIYFILTLFWGSFFGSI FMLS P  
FLPLMEVNP SWYRWINNRLVATWLTLPVALLETMFGVKVIITGDAFVP GERSVIIMNHRTRMDWM  
FLWNCLMRYSYLRLEKICLKASLKGVP GFGWAMQAAAYIFIHRKWKDDKSHFEDMIDYFCDIHEP  
LQLLIFPEGTDLTENSKSR SN AFAEKNGLQKYEYVLHPRTTGFTFVVDRLREGKNLDAVHDITVA  
YPHNIPQSEKHLLQGDFPREIH FHVHRYPIDTLPTS KEDLQLWCHKRWEEKEERLRSFYQGEKNF  
YFTGQSVIPPCKSELRVLVVKLLSILYWTLFSPAMCLLIYLYSLVKWYFIITIVIFVLQERIFGG  
LEIIELACYRLLHKQPHLNSKKNE

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-22

#### **Transmembrane domains:**

amino acids 44-63, 90-108, 354-377

**FIGURE 103**

CGGCTCGAGCGGCTCGAGTGAAGAGCCTCTCCACGGCTCCTGCGCCTGAGACAGCTGGCCTGACC  
 TCCAAATCATCCATCCACCCCTGCTGTCATCTGTTTTTCATAGTGTGAGATCAACCCACAGGAATA  
 TCCATGGCTTTTGTGCTCATTTTGGTTCTCAGTTTCTACGAGCTGGTGTGAGGACAGTGGCAAGT  
 CACTGGACCGGGCAAGTTTGTCCAGGCCCTTGGTGGGGGAGGACGCCGTGTTCTCCTGCTCCCTCT  
 TTCCTGAGACCAGTGCAGAGGCTATGGAAGTGCGGTTCTTCAGGAATCAGTTCCATGCTGTGGTC  
 CACCTCTACAGAGATGGGGAAGACTGGGAATCTAAGCAGATGCCACAGTATCGAGGGAGAACTGA  
 GTTTGTGAAGGACTCCATTGCAGGGGGGCGTGTCTCTCTAAGGCTAAAAACATCACTCCCTCGG  
 ACATCGGCCTGTATGGGTGCTGGTTTCAGTTCCCAGATTTACGATGAGGAGGCCACCTGGGAGCTG  
 CGGGTGGCAGCACTGGGCTCACTTCCTCTCATTTCCATCGTGGGATATGTTGACGGAGGTATCCA  
 GTTACTCTGCCTGTCTCAGGCTGGTTCCCCCAGCCACAGCCAAGTGGAAGGTCCACAAGGAC  
 AGGATTTGTCTTCAGACTCCAGAGCAAATGCAGATGGGTACAGCCTGTATGATGTGGAGATCTCC  
 ATTATAGTCCAGGAAAATGCTGGGAGCATATTGTGTTCCATCCACCTTGCTGAGCAGAGTCATGA  
 GGTGGAATCCAAGGTATTGATAGGAGAGACGTTTTTCCAGCCCTCACCTTGCGCCTGGCTTCTA  
 TTTTACTCGGGTACTCTGTGGTGCCCTGTGTGGTGTGTCATGGGGATGATAATTGTTTTCTTC  
 AAATCCAAAGGGAAAATCCAGGCGGAAGTGGACTGGAGAAGAAAGCACGGACAGGCAGAATTGAG  
 AGACGCCCGGAAACACGCAGTGGAGGTGACTCTGGATCCAGAGACGGCTCACCCGAAGCTCTGCG  
 TTTCTGATCTGAAAAGTGTAAACCATAGAAAAGCTCCCCAGGAGGTGCCTCACTCTGAGAAGAGA  
 TTTACAAGGAAGAGTGTGGTGGCTTCTCAGGGTTTCCAAGCAGGGAGACATTACTGGGAGGTGGA  
 CGTGGGACAAAATGTAGGGTGGTATGTGGGAGTGTGTGCGGATGACGTAGACAGGGGGAAGAACA  
 ATGTGACTTTGTCTCCCAACAATGGGTATTGGGTCTCAGACTGACAACAGAACATTTGTATTTT  
 ACATTCAATCCCCATTTTATCAGCCTCCCCCCCAGCACCCCTCCTACACGAGTAGGGGTCTTCCT  
 GGACTATGAGGGTGGGACCATCTCCTTCTTCAATACAAATGACCAGTCCCTTATTTATACCCTGC  
 TGACATGTCAGTTTGAAGGCTTGTTGAGACCCTATATCCAGCATGCGATGTATGACGAGGAAAAG  
 GGGACTCCCATATTCATATGTCCAGTGTCTTGGGGATTGAGACAGAGAAGACCCTGCTTAAAGGGC  
 CCCACACCACAGACCCAGACACAGCCAAGGGAGAGTGCTCCCGACAGGTGGCCCCAGCTTCCTCT  
 CCGGAGCCTGCGCACAGAGAGTCACGCCCCCACTCTCCTTTAGGGAGCTGAGGTTCTTCTGCCC  
 TGAGCCCTGCAGCAGCGGCAGTCACAGCTTCCAGATGAGGGGGGATTGGCCTGACCCTGTGGGAG  
 TCAGAAGCCATGGCTGCCCTGAAGTGGGGACGGAATAGACTCACATTAGGTTTAGTTTGTGAAAA  
 CTCCATCCAGCTAAGCGATCTTGAACAAGTCACAACCTCCCAGGCTCCTCATTTGCTAGTCACGG  
 ACAGTGATTCCTGCCTCACAGGTGAAGATTAAAGAGACAACGAATGTGAATCATGCTTGCAGGTT  
 TGAGGGGCACAGTGTGTTGCTAATGATGTGTTTTTATATTATACATTTTCCCACCATAAACTCTGTT  
 TGCTTATTCCACATTAATTTACTTTTCTCTATACCAAATCACCCATGGAATAGTTATTGAACACC  
 TGCTTTGTGAGGCTCAAAGAATAAAGAGGAGGTAGGATTTTTTCACTGATTCTATAAGCCCAGCAT  
 TACCTGATACCAAAACCAGGCAAAGAAAACAGAAGAAGAGGAAGGAAAACCTACAGGTCCATATCC  
 CTCATTAACACAGACACAAAAATTCTAAATAAAATTTTAACAAATTAACTAAACAATATATTTA  
 AAGATGATATATACTACTCAGTGTGGTTTGTCCACAAATGCAGAGTTGGTTTAATATTTAAAT  
 ATCAACCAGTGTAATTCAGCACATTAATAAAGTAAAAAAGAAAACCATAAAAAAAAAAAAAAA

## **FIGURE 104**

MAFVLILVLSFYELVSGQWQVTGPGKFVQALVGEDAVFSCSLFPETSAEAMEVRFFRNQFHAVVH  
LYRDGEDWESKQMPQYRGRTEFVKDSIAGGRVSLRLKNITPSDIGLYGCWFESSQIYDEEATWELR  
VAALGSLPLISIVGYVDGGIQLLCLSSGWFPQPTAKWKGPQGQDLSSDSRANADGYSLYDVEISI  
IVQENAGSILCSIHLAEQSHEVESKVLIGETFFQPSPWRLASILLGLLCGALCGVVMGMIIVFFK  
SKGKIQAELDWRRKHGQAEIRDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKAPQEVPHSEKRF  
TRKSVVASQGFQAGRHYWEVDVGQNVGWYVGVCRDDVDRGKNNVTLSPNNGYWVLRLTTEHLYFT  
FNPHFISLPPSTPPTRVGVFLDYEGGTISFFNTNDQSLIYTLLTCQFEGLLRPYIQHAMYDEEKG  
TPIFICPVSWG

**Signal peptide:**

amino acids 1-17

**Transmembrane domains:**

amino acids 131-150, 235-259

**FIGURE 105**

CCTTCACAGGACTCTTCATTGCTGGTTGGCA**ATG**ATGTATCGGCCAGATGTGGTGAGGGCTAGGA  
 AAAGAGTTTGTGGGAACCCTGGGTTATCGGCCCTCGTCATCTTCATATCCCTGATTGTCCTGGCA  
 GTGTGCATTGGACTCACTGTTTATTATGTGAGATATAATCAAAAGAAGACCTACAATTACTATAG  
 CACATTGTCAATTTACAACGACAACTATATGCTGAGTTTGGCAGAGAGGCTTCTAACAATTTTA  
 CAGAAATGAGCCAGAGACTTGAATCAATGGTGAAAAATGCATTTTATAAATCTCCATTAAGGGAA  
 GAATTTGTCAAGTCTCAGGTTATCAAGTTCAGTCAACAGAAGCATGGAGTGTGGCTCATATGCT  
 GTTGATTTGTAGATTTCACTCTACTGAGGATCCTGAAACTGTAGATAAAATTGTTCAACTTGTTT  
 TACATGAAAAGCTGCAAGATGCTGTAGGACCCCCCTAAAGTAGATCCTCACTCAGTTAAAATTTAA  
 AAAATCAACAAGACAGAAACAGACAGCTATCTAAACCATTGCTGCGGAACACGAAGAAGTAAAAC  
 TCTAGGTCAGAGTCTCAGGATCGTTGGTGGGACAGAAGTAGAAGAGGGTGAATGGCCCTGGCAGG  
 CTAGCCTGCAGTGGGATGGGAGTCATCGCTGTGGAGCAACCTTAATTAATGCCACATGGCTTGTTG  
 AGTGCTGCTCACTGTTTTACAACATATAAGAACCCTGCCAGATGGACTGCTTCCTTTGGAGTAAC  
 AATAAAACCTTCGAAAATGAAACGGGGTCTCCGGAGAATAATTGTCCATGAAAAATACAAACACC  
 CATCACATGACTATGATATTTCTCTTGCAGAGCTTTCTAGCCCTGTTCCCTACACAAATGCAGTA  
 CATAGAGTTTGTCTCCCTGATGCATCCTATGAGTTTCAACCAGGTGATGTGATGTTTGTGACAGG  
 ATTTGGAGCACTGAAAAATGATGGTTACAGTCAAAATCATCTTCGACAAGCACAGGTGACTCTCA  
 TAGACGCTACAACCTTGCAATGAACCTCAAGCTTACAATGACGCCATAACTCCTAGAATGTTATGT  
 GCTGGCTCCTTAGAAGGAAAAACAGATGCATGCCAGGGTGACTCTGGAGGACCACTGGTTAGTTC  
 AGATGCTAGAGATATCTGGTACCTTGCTGGAATAGTGAGCTGGGGAGATGAATGTGCGAAACCCA  
 ACAAGCCTGGTGTTTATACTAGAGTTACGGCCTTGCGGGACTGGATTACTTCAAAAACCTGGTATC  
**TAA**GAGACAAAAGCCTCATGGAACAGATAACATTTTTTTTTTGTTTTTTGGGTGTGGAGGCCATTT  
 TTAGAGATACAGAATTGGAGAAGACTTGCAAAACAGCTAGATTTGACTGATCTCAATAAACTGTT  
 TGCTTGATGCATGTATTTTCTTCCCAGCTCTGTTCCGCACGTAAGCATCCTGCTTCTGCCAGATC  
 AACTCTGTCATCTGTGAGCAATAGTTGAACTTTATGTACATAGAGAAATAGATAATACAATATT  
 ACATTACAGCCTGTATTCATTTGTTCTCTAGAAGTTTTGTGAGAATTTTGACTTGTTGACATAAA  
 TTTGTAATGCATATATACAATTTGAAGCACTCCTTTTCTTCAGTTCCTCAGCTCCTCTCATTTCA  
 GCAATATCCATTTTCAAGGTGCAGAACAAGGAGTGAAAGAAAATATAAGAAGAAAAAATCCCC  
 TACATTTTATTGGCACAGAAAAGTATTAGGTGTTTTTCTTAGTGGAATATTAGAAATGATCATAT  
 TCATTATGAAAGGTCAAGCAAAGACAGCAGAATACCAATCACTTCATCATTTAGGAAGTATGGGA  
 ACTAAGTTAAGGAAGTCCAGAAAGAAGCCAAGATATATCCTTATTTTCATTTCCAACAACACTACT  
 ATGATAAATGTGAAGAAGATTCTGTTTTTTGTGACCTATAATAATTATACAACTTCATGCAAT  
 GTACTTGTTCTAAGCAAATTAAAGCAAATATTTATTTAACATTGTTACTGAGGATGTCAACATAT  
 AACAAATAAAATATAAATCACCCA

## **FIGURE 106**

MMYRPDVVRARKRVCWEPWVIGLVIFISLIVLAVCIGLTVHYVRYNQKKTYNYYSTLSFTTDKLY  
AEFGREASNNFTEMSQRLESMVKNAFYKSPLREEFVKSQVIKFSQQKHGVLAHMLLICRFHSTED  
PETVDKIVQLVLHEKLQDAVGPPKVDPHSVKIKKINKTETDSYLNHCCGTRRSKTLGQSLRIVGG  
TEVEEGEWPWQASLQWDGSHRCGATLINATWLVSAAHCFTTYKNPARWTASFGVTIKPSKMKRGL  
RRIIVHEKYKHPSHDYDISLAELSSPVPTYNAVHRVCLPDASYEFQPGDVMFVTGFGALKNDGYS  
QNHLRQAQVTLIDATTCNEPQAYNDAITPRMLCAGSLEGKTDACQGDSGGPLVSSDARDIWYLAG  
IVSWGDECAKPNKPGVYTRVTALRDWITSKTGI

**Transmembrane domain:**

amino acids 21-40 (type II)

**FIGURE 107**

AGAGAAAGAAGCGTCTCCAGCTGAAGCCAATGCAGCCCTCCGGCTCTCCGCGAAGAAGTTCCTG  
 CCCCAGATGAGCCCCCGCCGTGCGTCCCCGACTATCCCCAGGCGGGCGTGGGGCACCGGGCCCAGC  
 GCCGACGATCGCTGCCGTTTTGCCCCTTGGGAGTAGGATGTGGTGAAAGGATGGGGCTTCTCCCTT  
 ACGGGGCTCACAA**ATG**GCCAGAGAAGATTCCGTGAAGTGTCTGCGCTGCCTGCTCTACGCCCTCAA  
 TCTGCTCTTTTTGGTTAATGTCCATCAGTGTGTTGGCAGTTTTCTGCTTGGATGAGGGACTACCTAA  
 ATAATGTTCTCACTTTAACTGCAGAAACGAGGGTAGAGGAAGCAGTCATTTTGACTTACTTTCCCT  
 GTGGTTCATCCGGTCATGATTGCTGTTTGCTGTTTCCTTATCATTGTGGGGATGTTAGGATATTG  
 TGAACGGTGAAAAGAAATCTGTTGCTTCTTGTCATGGTACTTTGGAAGTTTGCTTGTCAATTTCT  
 GTGTAGAACTGGCTTGTGGCGTTTGGACATATGAACAGGAACTTATGGTTCCAGTACAATGGTCA  
 GATATGGTCACTTTGAAAGCCAGGATGACAAATTATGGATTACCTAGATATCGGTGGCTTACTCA  
 TGCTTGGAAATTTTTTTTCAGAGAGAGTTTAAAGTGCTGTGGAGTAGTATATTTCACTGACTGGTTGG  
 AAATGACAGAGATGGACTGGCCCCCAGATTCCCTGCTGTGTTAGAGAATCCCAGGATGTTCCAAA  
 CAGGCCACCAGGAAGATCTCAGTGACCTTTATCAAGAGGGTTGTGGGAAGAAAATGTATTCCTT  
 TTTGAGAGGAACCAAACAACCTGCAGGTGCTGAGGTTTCTGGGAATCTCCATTGGGGTGACACAAA  
 TCCTGGCCATGATTCTCACCATTACTCTGCTCTGGGCTCTGTATTATGATAGAAGGGAGCCTGGG  
 ACAGACCAAATGATGTCTTGAAGAATGACAACTCTCAGCACCTGTCATGTCCCTCAGTAGAACT  
 GTTGAACCAAGCCTGTCAAGAATCTTTGAACACACATCCATGGCAAACAGCTTTAATACACACT  
 TTGAGATGGAGGAGTTAT**TAA**AAAAGAAATGTCACAGAAGAAAACCACAAACTTGTTTTATTGGACT  
 TGTGAATTTTTTGAGTACATACTATGTGTTTCAGAAATATGTAGAAATAAAAATGTTGCCATAAAA  
 TAACACCTAAGCATATACTATTCTATGCTTTAAATGAGGATGGAAAAGTTTCATGTCATAAGTC  
 ACCACCTGGACAATAATTGATGCCCTTAAATGCTGAAGACAGATGTCATACCCACTGTGTAGCC  
 TGTGTATGACTTTTACTGAACACAGTTATGTTTTGAGGCAGCATGGTTTGATTAGCATTTCCGCA  
 TCCATGCAAACGAGTCACATATGGTGGGACTGGAGCCATAGTAAAGGTTGATTACTTCTACCAA  
 CTAGTATATAAAGTACTAATTAAATGCTAACATAGGAAGTTAGAAAATACTAATAACTTTTATTA  
 CTCAGCGATCTATTCTTCTGATGCTAAATAAATTATATATCAGAAAACCTTCAATATTGGTGACT  
 ACCTAAATGTGATTTTTTGCTGGTTACTAAAATATTTCTTACCACTTAAAAGAGCAAGCTAACACAT  
 TGTCTTAAGCTGATCAGGGATTTTTTGTATATAAGTCTGTGTTAAATCTGTATAATTCAGTCGAT  
 TTCAGTTCTGATAATGTTAAGAATAACCATTATGAAAAGGAAAATTTGTCCTGTATAGCATCATT  
 ATTTTATAGCCTTTCCTGTTAATAAAGCTTTACTATTCTGTCCTGGGCTTATATTACACATATAAC  
 TGTTATTTAAATACTTAACCACTAATTTTGAAAATTACCAGTGTGATACATAGGAATCATTATTC  
 AGAATGTAGTCTGGTCTTTAGGAAGTATTAATAAGAAAATTTGCACATAACTTAGTTGATTCAGA  
 AAGGACTTGTATGCTGTTTTTCTCCCAAATGAAGACTCTTTTGGACACTAAACACTTTTTAAAAA  
 GCTTATCTTTGCCTTCTCCAAACAAGAAGCAATAGTCTCCAAGTCAATATAAATTTCTACAGAAAA  
 TAGTGTTCTTTTTCTCCAGAAAAATGCTTGTGAGAATCATTAACACATGTGACAATTTAGAGATT  
 CTTTGTTTTATTTCACTGATTAATATACTGTGGCAAATTACACAGATTATTAAATTTTTTTTACAA  
 GAGTATAGTATATTTATTTGAAATGGGAAAAGTGCATTTTACTGTATTTTGTGTATTTTGTGTTAT  
 TTCTCAGAATATGGAAAGAAAATTAATGTGTCAATAAATATTTTCTAGAGAGTAA

## **FIGURE 108**

MAREDSVKCLRCLLYALNLLFWLMSISVLAVSAWMRDYLNNVLTTLTAETRVEEAVILTYFPVVHP  
VMIAVCCFLIIVGMLGYCGTVKRNLLLLAWYFGSLLVIFCVELACGVWTYEQELMVPVQWSDMVT  
LKARMTNYGLPRYRWLTHAWNFFQREFKCCGVVYFTDWLEMTMDWPPDSCCVREFPGCSKQAHQ  
EDLSDLYQEGCGKKMYSFLRGTKQLQVLRFLGISIGVTQILAMILTITLLWALYYDRREPGTDQM  
MSLKNDNSQHLSCPSVELLKPSLSRIFEHTSMANSFNTHFEMEEL

**Signal peptide:**

amino acids 1-33

**Transmembrane domains:**

amino acids 12-35, 57-86, 94-114, 226-248



**FIGURE 109**

CCAAGGCCAGAGCTGTGGACACCTTATCCCACTCATCCTCATCCTCTTCCTCTGATAAAGCCCCCT  
 ACCAGTGCTGATAAAGTCTTTCTCGTGAGAGCCTAGAGGCCTTAAAAAAAAAAGTGCTTGAAAGA  
 GAAGGGGACAAAGGAACACCAGTATTAAGAGGATTTTCCAGTGTTTCTGGCAGTTGGTCCAGAAG  
**GATG**CCCTCCATTCTGCTTCTCACCTGCCTCTTCATCACAGGCACCTCCGTGTCACCCGTGGCCC  
 TAGATCCTTGTTCTGCTTACATCAGCCTGAATGAGCCCTGGAGGAACACTGACCACCAGTTGGAT  
 GAGTCTCAAGGTCCTCCTCTATGTGACAACCATGTGAATGGGGAGTGGTACCACTTCACGGGCAT  
 GCGGGGAGATGCCATGCCTACCTTCTGCATACCAGAAAACCACTGTGGAACCCACGCACCTGTCT  
 GGCTCAATGGCAGCCACCCCTAGAAAGGCGACGGCATTGTGCAACGCCAGGCTTGTGCCAGCTTC  
 AATGGGAACTGCTGTCTCTGGAACACCACGGTGGAAGTCAAGGCTTGCCCTGGAGGCTACTATGT  
 GTATCGTCTGACCAAGCCCAGCGTCTGCTTCCACGTCTACTGTGGTCATTTTTATGACATCTGCG  
 ACGAGGACTGCCATGGCAGCTGCTCAGATACCAGCGAGTGCACATGCGCTCCAGGAACTGTGCTA  
 GGCCCTGACAGGCAGACATGCTTTGATGAAAATGAATGTGAGCAAAACAACGGTGGCTGCAGTGA  
 GATCTGTGTGAACCTCAAAAACCTCCTACCGCTGTGAGTGTGGGGTTGGCCGTGTGCTAAGAAGTG  
 ATGGCAAGACTTGTGAAGACGTTGAAGGATGCCACAATAACAATGGTGGCTGCAGCCACTCTTGC  
 CTTGGATCTGAGAAAGGCTACCAAGTGTGAATGTCCCCGGGGCCTGGTGCTGTCTGAGGATAACCA  
 CACTTGCCAAGTCCCTGTGTTGTGCAAATCAAATGCCATTGAAGTGAACATCCCCAGGGAGCTGG  
 TTGGTGGCCTGGAGCTCTTCCTGACCAACACCTCCTGCCGAGGAGTGCCAACGGCACCCATGTC  
 AACATCCTCTTCTCTCTCAAGACATGTGGTACAGTGGTTCGATGTGGTGAATGACAAGATTGTGGC  
 CAGCAACCTCGTGACAGGTCTACCCAAGCAGACCCCGGGGAGCAGCGGGGACTTCATCATCCGAA  
 CCAGCAAGCTGCTGATCCCGGTGACCTGCGAGTTTCCACGCCTGTACACCATTTCTGAAGGATAC  
 GTTCCCAACCTTCGAAACTCCCCACTGGAAATCATGAGCCGAAATCATGGGATCTTCCCATTAC  
 TCTGGAGATCTTCAAGGACAATGAGTTTGAAGAGCCTTACCGGGAAGCTCTGCCCACCCTCAAGC  
 TTCGTGACTCCCTCTACTTTGGCATTGAGCCCGTGGTGCACGTGAGCGGCTTGGAAAGCTTGGTG  
 GAGAGCTGCTTTGCCACCCCCACCTCCAAGATCGACGAGGTCTGAAATACTACCTCATCCGGGA  
 TGGCTGTGTTTCAGATGACTCGGTAAAGCAGTACACATCCCGGGATCACCTAGCAAAGCACTTCC  
 AGGTCCCTGTCTTCAAGTTTGTGGGCAAAGACCACAAGGAAGTGTTTCTGCACTGCCGGGTCTT  
 GTCTGTGGAGTGTTGGACGAGCGTTCCCGCTGTGCCAGGGTTGCCACCGGCGAATGCGTCGTGG  
 GGCAGGAGGAGAGGACTCAGCCGGTCTACAGGGCCAGACGCTAACAGGCGGCCCGATCCGCATCG  
 ACTGGGAGGACT**TAG**TTCGTAGCCATACCTCGAGTCCCTGCATTGGACGGCTCTGCTCTTTGGAGC  
 TTCTCCCCCACCGCCCTCTAAGAACATCTGCCAACAGCTGGGTTCAGACTTCACACTGTGAGTT  
 CAGACTCCCAGCACCAACTCACTCTGATTCTGGTCCATTACAGTGGGCACAGGTCACAGCACTGCT  
 GAACAATGTGGCCTGGGTGGGGTTTCATCTTTCTAGGGTTGAAAACCTAACTGTCCACCCAGAAA  
 GACACTCACCCCATTTCCCTCATTTCTTTCCTACACTTAAATACCTCGTGTATGGTGAATCAGA  
 CCACAAAATCAGAAGCTGGGTATAATATTTCAAGTTACAAACCCTAGAAAAATTAAACAGTTACT  
 GAAATTATGACTTAAATACCCAATGACTCCTTAAATATGTAAATTATAGTTATACCTTGAAATTT  
 CAATTCAAATGCAGACTAATTATAGGGAATTTGGAAGTGTATCAATAAAACAGTATATAATTTT

## **FIGURE 110**

MPPFLLLTCLFITGTSVSPVALDPCSAYISLNEPWRNTDHQLDESQGPPLCDNHVNGEWYHFTGM  
 AGDAMPTFCIPENHCGTHAPVWLNGSHPLEGDGIVQRQACASFNGNCCLWNTTVEVKACPGGYV  
 YRLTKPSVCFHVYCGHFYDICDEDCHGSCSDTSECTCAPGTVLGPDRQTCFDENECEQNNGGCSE  
 ICVNLKNSYRCECGVGRVLRSDGKTCEDVEGCHNNNGGCSHSCLGSEKGYQCECPRGLVLSEDNH  
 TCQVPVLCKSNAIEVNI PRELVGGLELFLTNTSCRGVSNNGTHVNILFSLKTCGTVVDVNDKIVA  
 SNLVTGLPKQTPGSSGDFIIRTSKLLIPVTCEFPRLYTISEGYVPNLRNSPLEIMSRNHGIFPFT  
 LEIFKDNEFEOPYREALPTLKLRLDSLYFGIEPVVHVSGLESIVESCFATPTSKIDEVLKYYLIRD  
 GCVSDDSVKQYTSRDHLAKHFQVPVFKFVGKDHKEVFLHCRVLVCGVLDERSRCAQGCHRRMRRG  
 AGGEDSAGLQGQTLTGGPIRIDWED

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-16

#### **N-glycosylation sites.**

amino acids 89-93, 116-120, 259-263, 291-295, 299-303

#### **Tyrosine kinase phosphorylation sites.**

amino acids 411-418, 443-451

#### **N-myristoylation sites.**

amino acids 226-232, 233-239, 240-246, 252-258, 296-302, 300-306,  
 522-528, 531-537

#### **Aspartic acid and asparagine hydroxylation site.**

amino acids 197-209

#### **ZP domain proteins.**

amino acids 431-457

#### **Calcium-binding EGF-like proteins.**

amino acids 191-212, 232-253

**FIGURE 111**

GAGAGAGGCAGCAGCTTGCTCAGCGGACAAGGATGCTGGGCGTGAGGGACCAAGGCCTGCCCTGC  
 ACTCGGGCCTCCTCCAGCCAGTGCTGACCAGGGACTTCTGACCTGCTGGCCAGCCAGGACCTGTG  
 TGGGGAGGCCCTCCTGCTGCCTTGGGGTGACAATCTCAGCTCCAGGCTACAGGGAGACCGGGAGG  
 ATCACAGAGCCAGC**ATG**TTACAGGATCCTGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAA  
 CCCCTGCGCAAACCCCGTATCCCCATGGAGACCTTCAGAAAGGTGGGGATCCCCATCATCATAGC  
 ACTACTGAGCCTGGCGAGTATCATCATTGTGGTTGTCTCATCAAGGTGATTCTGGATAAATACT  
 ACTTCTCTGCGGGCAGCCTCTCCACTTCATCCCGAGGAAGCAGCTGTGTGACGGAGAGCTGGAC  
 TGTCCCTTGGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTCCCCGAAGGGCCTGCAGTGGCAGT  
 CCGCCTCTCCAAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCACAGGGAACTGGTTCTCTG  
 CCTGTTTCGACAACCTTCACAGAAGCTCTCGCTGAGACAGCCTGTAGGCAGATGGGCTACAGCAGA  
 GCTGTGGAGATTGGCCAGACCAGGATCTGGATGTTGTTGAAATCACAGAAAACAGCCAGGAGCT  
 TCGCATGCGGAACTCAAGTGGGCCCTGTCTCTCAGGCTCCCTGGTCTCCCTGCACTGTCTTGCCT  
 GTGGGAAGAGCCTGAAGACCCCCCGTGTGGTGGGTGGGGAGGAGGCCTCTGTGGATTCTTGGCCT  
 TGGCAGGTCAGCATCCAGTACGACAAACAGCACGTCTGTGGAGGGAGCATCCTGGACCCCCACTG  
 GGTCTCACGGCAGCCCACTGCTTCAGGAAACATAACCGATGTGTTCAACTGGAAGGTGCGGGCAG  
 GCTCAGACAAACTGGGCAGCTTCCCATCCCTGGCTGTGGCCAAGATCATCATCATTGAATTCAAC  
 CCCATGTACCCCAAAGACAATGACATCGCCCTCATGAAGCTGCAGTTCCCACTCACTTTCTCAGG  
 CACAGTCAGGCCCATCTGTCTGCCCTTCTTTGATGAGGAGCTCACTCCAGCCACCCCACTCTGGA  
 TCATTGGATGGGGCTTTACGAAGCAGAATGGAGGGAAGATGTCTGACATACTGCTGCAGGCGTCA  
 GTCCAGGTCATTGACAGCACACGGTGCAATGCAGACGATGCGTACCAGGGGGAAGTCACCGAGAA  
 GATGATGTGTGCAGGCATCCCGGAAGGGGGTGTGGACACCTGCCAGGGTGACAGTGGTGGGCCCC  
 TGATGTACCAATCTGACCAGTGGCATGTGGTGGGCATCGTTAGCTGGGGCTATGGCTGCGGGGGC  
 CCGAGCACCCCAAGAGTATACACCAAGGTCTCAGCCTATCTCAACTGGATCTACAATGTCTGGAA  
 GGCTGAGCTG**TAA**TGCTGCTGCCCCCTTTCAGAGTGTGGGAGCCGCTTCCTTCCTGCCCCGCCCCAC  
 CTGGGGATCCCCCAAAGTCAGACACAGAGCAAGAGTCCCCTTGGGTACACCCCTCTGCCCCACAGC  
 CTCAGCATTTCTTGGAGCAGCAAAGGGCCTCAATTCCCTGTAAGAGACCCTCGCAGCCCAGAGGCG  
 CCCAGAGGAAGTCAGCAGCCCTAGCTCGGCCACACTTGGTGCTCCCAGCATCCCAGGGAGAGACA  
 CAGCCCACTGAACAAGGTCTCAGGGGTATTGCTAAGCCAAGAAGGAACCTTCCCACACTACTGAA  
 TGGAAGCAGGCTGTCTTGTAAGCCAGATCACTGTGGGCTGGAGAGGAGAAGGAAAGGGTCTG  
 CGCCAGCCCTGTCCGTCTTACCCATCCCCAAGCCTACTAGAGCAAGAAACCAGTTGTAATATAA  
 AATGCACTGCCCTACTGTTGGTATGACTACCGTTACCTACTGTTGTCATTGTTATTACAGCTATG  
 GCCACTATTATTAAAGAGCTGTGTAACATCTCTGGCAAAAAAAAAAAAA

## **FIGURE 112**

MLQDPDSDQPLNSLDVKPLRKPRIPMETFRKVGIPIIIIALLSLASIIIVVVLIKVILDKYYFLCG  
QPLHFIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAVAVRLSKDRSTLQVLDSATGNWFSACFDN  
FTEALAETACRQMGYSRAVEIGPDQDLDVVEITENSQELMRNSSGPCLSGSLVSLHCLACGKSL  
KTPRVVGEEASVDSWPWQVSIQYDKQHVCGGSILDPHWVLTAAHCFRKHTDVFNWKVRAGSDKL  
GSFPSLAVAKIIIIIEFNPMYPKDNDIALMKLQFPLTFSGTVRPICLPFFDEELTPATPLWIIGWG  
FTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMCAGIPEGGVDTCCQGDSSGGLMYQS  
DQWHVVGIVSWGYGCGGPSTPGVYTKVSAYLNWIYNVWKAEL

### **Transmembrane domain:**

amino acids 32-53 (typeII)

**FIGURE 113**

GGCTGGACTGGAACCTCCTGGTCCCAAGTGATCCACCCGCCTCAGCCTCCCAAGGTGCTGTGATTA  
TAGGTGTAAGCCACCGTGTCTGGCCTCTGAACAACCTTTTTTCAGCAACTAAAAAGCCACAGGAGT  
TGAACCTGCTAGGATTCTGACTATGCTGTGGTGGCTAGTGCTCCTACTCCTACCTACATTAAATC  
TGTTTTTTGTTCTCTTGTAACCTAGCCTTTACCTTCCTAACACAGAGGATCTGTCACTGTGGCTCT  
GGCCCAAACCTGACCTTCACTCTGGAACGAGAACAGAGGTTTCTACCCACACCGTCCCCTCGAAG  
CCGGGGACAGCCTCACCTTGCTGGCCTCTCGCTGGAGCAGTGCCCTCACCAACTGTCTCACGTCT  
GGAGGCACTGACTCGGGCAGTGCAGGTAGCTGAGCCTCTTGGTAGCTGCGGCTTTCAAGGTGGGC  
CTTGCCCTGGCCGTAGAAGGGATTGACAAGCCCCGAAGATTTTCATAGGCGATGGCTCCCCTGCCC  
AGGCATCAGCCTTGCTGTAGTCAATCACTGCCCTGGGGCCAGGACGGGCGTGGACACCTGCTCA  
GAAGCAGTGGGTGAGACATCACGCTGCCCGCCCATCTAACCTTTTCATGTCCTGCACATCACCTG  
ATCCATGGGCTAATCTGAACTCTGTCCCAAGGAACCCAGAGCTTGAGTGAGCTGTGGCTCAGACC  
CAGAAGGGGTCTGCTTAGACCACCTGGTTTATGTGACAGGACTTGCATTCTCCTGGAACATGAGG  
GAACGCCGGAGGAAAGCAAAGTGGCAGGGAAGGAACTTGTGCCAAATTATGGGTGAGAAAAGATG  
GAGGTGTTGGGTATCACAAGGCATCGAGTCTCCTGCATTTCAGTGGACATGTGGGGGAAGGGCTG  
CCGATGGCGCATGACACACTCGGGACTCACCTCTGGGGCCATCAGACAGCCGTTTCCGCCCCGAT  
CCACGTACCAGCTGCTGAAGGGCAACTGCAGGCCGATGCTCTCATCAGCCAGGCAGCAGCCAAAA  
TCTGCGATCACCAGCCAGGGGCAGCCGTCTGGGAAGGAGCAAGCAAAGTGACCATTTCTCCTCCC  
CTCCTTCCCTCTGAGAGGCCCTCCTATGTCCCTACTAAAGCCACCAGCAAGACATAGCTGACAGG  
GGCTAATGGCTCAGTGTTGGCCCAGGAGGTCAGCAAGGCCTGAGAGCTGATCAGAAGGGCCTGCT  
GTGCGAACACGGAAATGCCTCCAGTAAGCACAGGCTGCAAAATCCCCAGGCAAAGGACTGTGTGG  
CTCAATTTAAATCATGTTCTAGTAATTGGAGCTGTCCCCAAGACCAAAGGAGCTAGAGCTTGGTT  
CAAATGATCTCCAAGGGCCCTTATACCCAGGAGACTTTGATTTGAATTTGAAACCCCAAATCCA  
AACCTAAGAACCAGGTGCATTAAGAATCAGTTATTGCCGGGTGTGGTGGCCTGTAATGCCAACAT  
TTTGGGAGGCCGAGGCGGGTAGATCACCTGAGGTCAGGAGTTCAAGACCAGCCTGGCCAACATGG  
TGAAACCCCTGTCTCTACTAAAAATACAAAAAACTAGCCAGGCATGGTGGTGTGTGCCTGTATC  
CCAGCTACTCGGGAGGCTGAGACAGGAGAATTACTTGAACCTGGGAGGTGAAGGAGGCTGAGACA  
GGAGAATCACTTCAGCCTGAGCAACACAGCGAGACTCTGTCTCAGAAAAATAAAAAAGAATTA  
TGGTTATTTGTAA

114/168

## **FIGURE 114**

MLWWLVLLLLPTLKSVFCSLVTSLYLPNTEDLSLWLWPKPDLHSGTRTEVSTHTVPSKPGTASPC  
WPLAGAVPSPTVSRLEALTRAVQVAEPLGSCGFQGGPCPGRRRD

**Signal peptide:**

amino acids 1-15

**FIGURE 115**

CAGCAGTGGTCTCTCAGTCCTCTCAAAGCAAGGAAAGAGTACTGTGTGCTGAGAGACCATGGCAA  
AGAATCCTCCAGAGAATTGTGAAGACTGTCACATTCTAAATGCAGAAGCTTTTAAATCCAAGAAA  
ATATGTAAATCACTTAAGATTTGTGGACTGGTGTGGTATCCTGGCCCTAACTCTAATTGTCCT  
GTTTTGGGGGAGCAAGCACTTCTGGCCGGAGGTACCCAAAAAGCCTATGACATGGAGCACACTT  
TCTACAGCAATGGAGAGAAGAAGAAGATTTACATGGAAATTGATCCTGTGACCAGAACTGAAATA  
TTCAGAAGCGGAAATGGCACTGATGAAACATTGGAAGTGCACGACTTTAAAAACGGATACACTGG  
CATCTACTTCGTGGGTCTTCAAAAATGTTTTATCAAACTCAGATTAAAGTGATTCCTGAATTTT  
CTGAACCAGAAGAGGAAATAGATGAGAATGAAGAAATTACCACAACCTTTCTTTGAACAGTCAGTG  
ATTTGGGTCCCAGCAGAAAAGCCTATTGAAAACCGAGATTTTCTTAAAAATTCCAAAATTCTGGA  
GATTTGTGATAACGTGACCATGTATTGGATCAATCCCACTCTAATATCAGTTTCTGAGTTACAAG  
ACTTTGAGGAGGAGGGAGAAGATCTTCACTTTCCTGCCAACGAAAAAAAGGGATTGAACAAAAT  
GAACAGTGGGTGGTCCCTCAAGTGAAAGTAGAGAAGACCCGTCACGCCAGACAAGCAAGTGAGGA  
AGAACTTCCAATAAATGACTATACTGAAAATGGAATAGAATTTGATCCCATGCTGGATGAGAGAG  
GTTATTGTTGTATTTACTGCCGTGAGGCAACCGCTATTGCCGCCGCGTCTGTGAACCTTTACTA  
GGCTACTACCCATATCCATACTGCTACCAAGGAGGACGAGTCATCTGTCGTGTCATCATGCCTTG  
TAACTGGTGGGTGGCCCGCATGCTGGGGAGGGTCTAATAGGAGGTTTGAGCTCAAATGCTTAAAC  
TGCTGGCAACATATAATAAATGCATGCTATTCAATGAATTTCTGCCTATGAGGCATCTGGCCCCT  
GGTAGCCAGCTCTCCAGAATTACTTGTAGGTAATTCCTCTCTTCATGTTCTAATAAACTTCTACA  
TTATCACCAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 116**

MAKNPPENCEDCHILNAEAFKSKKICKSLKICGLVFGILALTIVLFWGSKHFWPEVPPKKAYDME  
HTFYSNGEKKKIYMEIDPVT RTEIFRSGNGTDETLEVHDFKNGYTG IYFVGLQKCFIKTQIKVIP  
EFSEPEEEIDENEEITTTFFEQSVI WVPAEKPIENRDFLKN SKILEICDNVTMYWINPTLISVSE  
LQDFEEEGEDLHF PANEEKGIEQNEQWVVPQVKVEKTRHARQASEEELPINDYTENGIEFDPMLD  
ERGYCC IYCRRGNRYCRRVCEPLLGYYPYPYCYQGGRVICRVIMPCNWWVARMLGRV

**Important features of the protein:**

**Signal peptide:**

amino acids 1-40

**Transmembrane domain:**

amino acids 25-47 (type II)

**N-glycosylation sites.**

amino acids 94-97, 180-183

**Glycosaminoglycan attachment sites.**

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-242

**N-myristoylation sites.**

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

**Microbodies C-terminal targeting signal.**

amino acids 315-317

**Cytochrome c family heme-binding site signature.**

amino acids 9-14



**FIGURE 117**

GAGCTCCCCTCAGGAGCGCGTTAGCTTCACACCTTCGGCAGCAGGAGGGCGGCAGCTTCTCGCAG  
GCGGCAGGGCGGGCGGCCAGGATC**ATGT**TCCACCACCACATGCCAAGTGGTGGCGTTCCCTCCTGTC  
CATCCTGGGGCTGGCCGGCTGCATCGCGGCCACCGGGATGGACATGTGGAGCACCCAGGACCTGT  
ACGACAACCCCGTCACCTCCGTGTTCCAGTACGAAGGGCTCTGGAGGAGCTGCGTGAGGCAGAGT  
TCAGGCTTCACCGAATGCAGGCCCTATTTACCATCCTGGGACTTCCAGCCATGCTGCAGGCAGT  
GCGAGCCCTGATGATCGTAGGCATCGTCCTGGGTGCCATTGGCCTCCTGGTATCCATCTTTGCCC  
TGAAATGCATCCGCATTGGCAGCATGGAGGACTCTGCCAAAGCCAACATGACACTGACCTCCGGG  
ATCATGTTTATTGTCTCAGGTCTTTGTGCAATTGCTGGAGTGTCTGTGTTTGCCAACATGCTGGT  
GACTAACTTCTGGATGTCCACAGCTAACATGTACACCGGCATGGGTGGGATGGTGCAGACTGTTT  
AGACCAGGTACACATTTGGTGCGGCTCTGTTCGTGGGCTGGGTGCGTGGAGGCCTCACACTAATT  
GGGGGTGTGATGATGTGCATCGCCTGCCGGGGCCTGGCACCAGAAGAAACCAACTACAAAGCCGT  
TTCTTATCATGCCTCAGGCCACAGTGTTGCCTACAAGCCTGGAGGCTTCAAGGCCAGCACTGGCT  
TTGGGTCCAACACCAAAAAACAAGAAGATATACGATGGAGGTGCCCCGACAGAGGACGAGGTACAA  
TCTTATCCTTCCAAGCACGACTATGTG**TAAT**GTCTCTAAGACCTCTCAGCACGGGCGGAAGAACT  
CCCGGAGAGCTCACCCAAAAACAAGGAGATCCCATCTAGATTTCTTCTTGCTTTTGACTCACAG  
CTGGAAGTTAGAAAAGCCTCGATTTTCATCTTTGGAGAGGCCAAATGGTCTTAGCCTCAGTCTCTG  
TCTCTAAATATTCCACCATAAAACAGCTGAGTTATTTATGAATTAGAGGCTATAGCTCACATTTT  
CAATCCTCTATTTCTTTTTTTTAAATATAACTTTCTACTCTGATGAGAGAATGTGGTTTTAATCTC  
TCTCTCACATTTTGATGATTTAGACAGACTCCCCCTCTTCCTCCTAGTCAATAAACCCATTGATG  
ATCTATTTCCCAGCTTATCCCCAAGAAAACTTTTGAAAGGAAAGAGTAGACCCAAAGATGTTATT  
TTCTGCTGTTTGAATTTTGTCTCCCCACCCCCAAGTTGGCTAGTAATAAACACTTACTGAAGAAG  
AAGCAATAAGAGAAAGATATTTGTAATCTCTCCAGCCCATGATCTCGGTTTTCTTACACTGTGAT  
CTTAAAAGTTACCAAACCAAAGTCATTTTCAGTTTGAGGCAACCAAACCTTTCTACTGCTGTTGA  
CATCTTCTTATTACAGCAACACCATTCTAGGAGTTTCTGAGCTCTCCACTGGAGTCCCTCTTTCT  
GTCGCGGGTCAGAAATTGTCCCTAGATGAATGAGAAAATTATTTTTTTTAAATTAAGTCCTAAAT  
ATAGTTAAAATAAATAATGTTTTTAGTAAAATGATACACTATCTCTGTGAAATAGCCTCACCCCTA  
CATGTGGATAGAAGGAAATGAAAAAATAATTGCTTTGACATTGTCTATATGGTACTTTGTAAAGT  
CATGCTTAAGTACAAATTCCATGAAAAGCTCACACCTGTAATCCTAGCACTTTGGGAGGCTGAGG  
AGGAAGGATCACTTGAAGCCAGAAGTTCGAGACTAGCCTGGGCAACATGGAGAAGCCCTGTCTCT  
ACAAAATACAGAGAGAAAAAATCAGCCAGTCATGGTGGCATAACCTGTAGTCCCAGCATTTCCGG  
GAGGCTGAGGTGGGAGGATCACTTGAAGCCAGGGAGGTTGGGGCTGCAGTGAGCCATGATCACAC  
CACTGCACTCCAGCCAGGTGACATAGCGAGATCCTGTCTAAAAAAATAAAAAATAAATAATGGAA  
CACAGCAAGTCCTAGGAAGTAGGTTAAAACATAATTCTTTAA

## **FIGURE 118**

MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLRSCVRQSSGFTECRP  
YFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLTSGIMFIVSGL  
CAIAGVSVFANMLVTNFWMSTANMYTGMGGMVQTVQTRYTFGAALFVGWVAGGLTLIGGVMMCIA  
CRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFGSNTKNKKIYDGGARTEDEVQSYPSKHDY  
V

**Signal peptide:**

amino acids 1-23

**Transmembrane domains:**

amino acids 81-100, 121-141, 173-194

**FIGURE 119**

GGAAAACTGTTCTCTTCTGTGGCACAGAGAACCCTGCTTCAAAGCAGAAGTAGCAGTTCCGGAG  
 TCCAGCTGGCTAAAACTCATCCCAGAGGATA**ATG**GCAACCCATGCCTTAGAAATCGCTGGGCTGT  
 TTCTTGGTGGTGTGGAATGGTGGGCACAGTGGCTGTCACTGTGCCTCAGTGGAGAGTGTCTG  
 GCCTTCATTGAAAACAACATCGTGGTTTTTTGAAAACCTTCTGGGAAGGACTGTGGATGAATTGCGT  
 GAGGCAGGCTAACATCAGGATGCAGTGCAAAATCTATGATTCCCTGCTGGCTCTTTCTCCGGACC  
 TACAGGCAGCCAGAGGACTGATGTGTGCTGCTTCCGTGATGTCCTTCTTGGCTTTCATGATGGCC  
 ATCCTTGGCATGAAATGCACCAGGTGCACGGGGGACAATGAGAAGGTGAAGGCTCACATTCTGCT  
 GACGGCTGGAATCATCTTCATCATCACGGGCATGGTGGTGTCTCATCCCTGTGAGCTGGGTGGCA  
 ATGCCATCATCAGAGATTTCTATAACTCAATAGTGAATGTTGCCCAAAAACGTGAGCTTGGAGAA  
 GCTCTCTACTTAGGATGGACCACGGCACTGGTGGTGTGATTGTTGGAGGAGCTCTGTTCTGCTGCGT  
 TTTTGTGTGCAACGAAAAGAGCAGTAGCTACAGATACTCGATACCTTCCCATCGCACAAACCCAAA  
 AAAGTTATCACACCGGAAAGAAGTCACCGAGCGTCTACTCCAGAAGTCAGTATGTG**TAG**TTGTGT  
 ATGTTTTTTTTAACTTTTACTATAAAGCCATGCAAATGACAAAAATCTATATTACTTTCTCAAAATG  
 GACCCCAAAGAACTTTGATTTACTGTTCTTAACTGCCTAATCTTAATTACAGGAAGTGTGCATC  
 AGCTATTTATGATTCTATAAGCTATTTTCAAGCAGAATGAGATATTAAACCCAATGCTTTGATTGTT  
 CTAGAAAGTATAGTAATTTGTTTTCTAAGGTGGTTCAAGCATCTACTCTTTTTATCATTTACTTC  
 AAAATGACATTGCTAAAGACTGCATTATTTTACTACTGTAATTTCTCCACGACATAGCATTATGT  
 ACATAGATGAGTGTAACATTTATATCTCACATAGAGACATGCTTATATGGTTTTATTTAAAATGA  
 AATGCCAGTCCATTACACTGAATAAATAGAACTCAACTATTGCTTTTCAGGGAAATCATGGATAG  
 GGTGAAGAAGGTACTATTAATTGTTTAAAAACAGCTTAGGGATTAATGTCCTCCATTTATAAT  
 GAAGATTAAAATGAAGGCTTTAATCAGCATTGTAAAGGAAATTGAATGGCTTTCTGATATGCTGT  
 TTTTAGCCTAGGAGTTAGAAATCCTAACTTCTTTATCCTCTTCTCCCAGAGGCTTTTTTTTTCT  
 TGTGTATTAAATTAACATTTTTTAAACGCAGATATTTTGTCAAGGGGCTTTCATTCAAACCTGCT  
 TTTCCAGGGCTATACTCAGAAGAAAGATAAAAGTGTGATCTAAGAAAAAGTGATGGTTTTAGGAA  
 AGTGAAAATATTTTTGTTTTGTATTTGAAGAAGAATGATGCATTTTGACAAGAAATCATATATG  
 TATGGATATATTTTAATAAGTATTTGAGTACAGACTTTGAGGTTTCATCAATATAAATAAAAGAG  
 CAGAAAAATATGTCTTGGTTTTTCATTTGCTTACCAAAAAACAACAACAAAAAAGTTGTCCTTT  
 GAGAACTTCACCTGCTCCTATGTGGGTACCTGAGTCAAATTTGTCATTTTTGTTCTGTGAAAAAT  
 AAATTTCTTCTTGTACCATTTCTGTTTAGTTTTACTAAAATCTGTAAATACTGTATTTTTCTGT  
 TTATTCCAAATTTGATGAACTGACAATCCAATTTGAAAGTTTGTGTCGACGTCTGTCTAGCTTA  
 AATGAATGTGTTCTATTTGCTTTATACATTTATATTAATAAATTGTACATTTTTCTAATT

120/168

## **FIGURE 120**

MATHALEIAGLFLGGVGMVGTVAVTVMPPQWRVSAFIENNIVVFENFWEGLWMNCVRQANIRMQCK  
IYDSLLALSPDLQAARGLMCAASVMSFLAFMMAILGMKCTRCTGDNEKVKAHILLTAGIIFIITG  
MVLIPVSWVANAIIRDFYNSIVNVAQKRELGEALYLGWTTALVLIVGGALFCCVFCCNEKSSSY  
RYSIPSHRTTQKSYHTGKKSPSVYSRSQYV

### **Signal peptide:**

amino acids 1-17

### **Transmembrane domains:**

amino acids 82-101, 118-145, 164-188

**FIGURE 121**

GGAGAGAGGCGCGCGGGTGAAAGGCGCATTGATGCAGCCTGCGGCGGCCTCGGAGCGCGGCGGAG  
CCAGACGCTGACCACGTTCTCTCCTCGGTCTCCTCCGCCTCCAGCTCCGCGCTGCCCCGGCAGCC  
GGGAGCCATGCGACCCCAGGGCCCCGCGCCTCCCCGCAGCGGCTCCGCGGCCTCCTGCTGCTCC  
TGCTGCTGCAGCTGCCCCGCGCCGTCGAGCGCCTCTGAGATCCCCAAGGGGAAGCAAAAGGCGCAG  
CTCCGGCAGAGGGAGGTGGTGGACCTGTATAATGGAATGTGCTTACAAGGGCCAGCAGGAGTGCC  
TGGTCGAGACGGGAGCCCTGGGGCCAATGTTATTCCGGGTACACCTGGGATCCCAGGTCGGGATG  
GATTCAAAGGAGAAAAGGGGGAATGTCTGAGGGAAAGCTTTGAGGAGTCCTGGACACCCAACTAC  
AAGCAGTGTTTCATGGAGTTCATTGAATTATGGCATAGATCTTGGGAAAATTGCGGAGTGACATT  
TACAAAGATGCGTTCAAATAGTGCTCTAAGAGTTTTGTTTCAGTGGCTCACTTCGGCTAAAATGCA  
GAAATGCATGCTGTCAGCGTTGGTATTTACATTCGAATGGAGCTGAATGTTTCAGGACCTCTTCCC  
ATTGAAGCTATAATTTATTTGGACCAAGGAAGCCCTGAAATGAATTCAACAATTAATATTCATCG  
CACTTCTTCTGTGGAAGGACTTTGTGAAGGAATTGGTGCTGGATTAGTGGATGTTGCTATCTGGG  
TTGGCACTTGTTTCAGATTACCCAAAAGGAGATGCTTCTACTGGATGGAATTCAGTTTCTCGCATC  
ATTATTGAAGAACTACCAAAATAAATGCTTTAATTTTCATTTGCTACCTCTTTTTTTATTATGCC  
TTGGAATGGTTCACTTAAATGACATTTTAAATAAGTTTATGTATACATCTGAATGAAAAGCAAAG  
CTAAATATGTTTACAGACCAAAGTGTGATTTACACTGTTTTTAAATCTAGCATTATTCATTTTG  
CTTCAATCAAAGTGGTTTCAATATTTTTTTTTAGTTGGTTAGAATACTTTCTTCATAGTCACATT  
CTCTCAACCTATAATTTGGAATATTGTTGTGGTCTTTTGTTTTTTCTCTTAGTATAGCATTTTTTA  
AAAAAATATAAAAGCTACCAATCTTTGTACAATTTGTAAATGTTAAGAATTTTTTTTTATATCTGT  
TAAATAAAAATTATTTCCAACA

122/168

## **FIGURE 122**

MRPQGPAASPQRLRGLLLLLLLQLPAPSSASEIPKGKQKAQLRQREVVDLYNGMCLQGPAGVPGR  
DGSPGANVIPGTPGIPGRDGFKEKGECLRESFEESWTPNYKQCSWSSLNYGIDLGKIAECTFTK  
MRSNSALRVLFSGSLRLKCRNACCQRWYFTFNGAECGGLPIEAIITYLDQGSPENNSTINIHRTS  
SVEGLCEGIGAGLVDVAIWVGTCSDYPKGDASTGWNSVSRIIEELPK

**Signal peptide:**

amino acids 1-30

**Transmembrane domain:**

amino acids 195-217

**FIGURE 123**

GCTGAGCGTGTGCGCGGTACGGGGCTCTCCTGCCTTCTGGGCTCCAACGCAGCTCTGTGGCTGAA  
CTGGGTGCTCATCACGGGAAGTCTGGGCTATGGAATACAGATGTGGCAGCTCAGGTAGCCCCAA  
ATTGCCTGGAAGAATACATCATGTTTTTCGATAAGAAGAAATTGTAGGATCCAGTTTTTTTTTTA  
ACCGCCCCCTCCCCACCCCCCAAAAAAAGTGTAAAGATGCAAAAACGTAATATCCATGAAGATCC  
TATTACCTAGGAAGATTTTGATGTTTTGCTGCGAATGCGGTGTTGGGATTTATTTGTTCTTGGAG  
TGTTCTGCGTGGCTGGCAAAGAATAATGTTCCAAAATCGGTCCATCTCCCAAGGGGTCCAATTTT  
TCTTCCTGGGTGTCAGCGAGCCCTGACTCACTACAGTGCAGCTGACAGGGGCTGTCATGCAACTG  
GCCCCTAAGCCAAAGCAAAAGACCTAAGGACGACCTTTGAACAATACAAAGG**ATG**GGTTTCAATG  
TAATTAGGCTACTGAGCGGATCAGCTGTAGCACTGGTTATAGCCCCACTGTCTTACTGACAATG  
CTTTCTTCTGCCGAACGAGGATGCCCTAAGGGCTGTAGGTGTGAAGGCAAAATGGTATATTGTGA  
ATCTCAGAAATTACAGGAGATACCCTCAAGTATATCTGCTGGTTGCTTAGGTTTGTCCCTTCGCT  
ATAACAGCCTTCAAAAACCTAAGTATAATCAATTTAAAGGGCTCAACCAGCTCACCTGGCTATAC  
CTTGACCATAACCATATCAGCAATATTGACGAAAATGCTTTTAATGGAATACGCAGACTCAAAGA  
GCTGATTCTTAGTTCCAATAGAATCTCCTATTTTCTTAACAATACCTTCAGACCTGTGACAAATT  
TACGGAACTTGGATCTGTCCTATAATCAGCTGCATTCTCTGGGATCTGAACAGTTTCGGGGCTTG  
CGGAAGCTGCTGAGTTTACATTTACGGTCTAACTCCCTGAGAACCATCCCTGTGCGAATATTCCA  
AGACTGCCGCAACCTGGAACCTTTGGACCTGGGATATAACCGGATCCGAAGTTTAGCCAGGAATG  
TCTTTGCTGGCATGATCAGACTCAAAGAACTTCACTGGAGCACAATCAATTTTCCAAGCTCAAC  
CTGGCCCTTTTTTCCAAGGTTGGTCAGCCTTCAGAACCTTTACTTGCAGTGGAATAAAATCAGTGT  
CATAGGACAGACCATGTCCTGGACCTGGAGCTCCTTACAAAGGCTTGATTTATCAGGCAATGAGA  
TCGAAGCTTTCAGTGGACCCAGTGTTTTCCAGTGTGTCCCGAATCTGCAGCGCCTCAACCTGGAT  
TCCAACAAGCTCACATTTATTGGTCAAGAGATTTTGGATTCTTGGATATCCCTCAATGACATCAG  
TCTTGCTGGGAATATATGGGAATGCAGCAGAAATATTTGCTCCCTTGTAAGTGGCTGAAAAGTT  
TTAAAGGTCTAAGGGAGAATACAATTATCTGTGCCAGTCCCAAAGAGCTGCAAGGAGTAAATGTG  
ATCGATGCAGTGAAGAACTACAGCATCTGTGGCAAAGTACTACAGAGAGGTTTGATCTGGCCAG  
GGCTCTCCCAAAGCCGACGTTTAAGCCCAAGCTCCCCAGGCCGAAGCATGAGAGCAAACCCCTT  
TGCCCCCGACGGTGGGAGCCACAGAGCCCGGCCAGAGACCGATGCTGACGCCGAGCACATCTCT  
TTCCATAAAATCATCGCGGGCAGCGTGGCGCTTTTTCCTGTCCGTGCTCGTCATCCTGCTGGTTAT  
CTACGTGTCATGGAAGCGGTACCCTGCGAGCATGAAGCAGCTGCAGCAGCGCTCCCTCATGCGAA  
GGCACAGGAAAAAGAAAAGACAGTCCCTAAAGCAAATGACTCCAGCACCCAGGAATTTTATGTA  
GATTATAAACCCACCAACACGGAGACCAGCGAGATGCTGCTGAATGGGACGGGACCCTGCACCTA  
TAACAAATCGGGCTCCAGGGAGTGTGAGGTAT**TGA**ACCATTGTGATAAAAAGAGCTCTTAAAAGCT  
GGGAAATAAGTGGTGCTTTATTGAACTCTGGTGACTATCAAGGGAACGCGATGCCCCCCTCCCC  
TTCCCTCTCCCTCTCACTTTGGTGGCAAGATCCTTTCCTTGTCCGTTTTAGTGCATTTCATAATACT  
GGTCATTTTCCCTCTCATACATAATCAACCCATTGAAATTTAAATACCACAATCAATGTGAAGCTT  
GAACTCCGGTTTAATATAATACCTATTGTATAAGACCCTTTACTGATTCCATTAATGTGCGATTT  
GTTTTAAGATAAAAACCTTCTTTCATAGGTAAAAA

## **FIGURE 124**

MGFNVIRLLSGSAVALVIAPT VLLTMLSSAERGCPKGCRCGKMVYCESQKLQEIPSSISAGCLG  
LSLRYNLSLQKLKYNQFKGLNQLTWLYLDHNNHISNIDENAFNGIRRLKELILSSNRISYFLNNTFR  
PVTNLRNLDLSYNQLHSLGSEQFRGLRKLLSLHLRSNSLRTIPVRI FQDCRNLELLDLGYNRIRS  
LARNVFAGMIRLKEHLEHNQFSKLNALFPRLVSLQONLYLQWNKISVIGQTMSTWSSLQRLDL  
SGNEIEAFSGPSVFQCVPNLQRLNLD SNKLTFIGQEILDSWISLNDISLAGNIWECSRNICSLVN  
WLKSFKGLRENTIICASPKELOGVNVIDAVKNYSICGKSTTERFDLARALPKPTFKPKLPRPKHE  
SKPPLPPTVGATEPGPETDADAEHISFHKIIAGSVALFLSVLVILLVIYVSWKRYPASMKQLQQR  
SLMRRHRKKKRQSLKQMT PSTQEFYVDYKPTNTETSEM L L NGTGPCTYNKSGSRECEV

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-33

#### **Transmembrane domain:**

amino acids 420-442

#### **N-glycosylation sites.**

amino acids 126-129, 357-360, 496-499, 504-507

#### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 465-468

#### **Tyrosine kinase phosphorylation site.**

amino acids 136-142

#### **N-myristoylation sites.**

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512



**FIGURE 125**

CCGTTATCGTCTTGCGCTACTGCTGA**ATG**TCCGTCCCGGAGGAGGAGGAGAGGCTTTTGCCGCTG  
ACCCAGAGATGGCCCCGAGCGAGCAAATTCCTACTGTCCGGCTGCGCGGCTACCGTGGCCGAGCT  
AGCAACCTTTCCCCTGGATCTCACAAAACCTCGACTCCAAATGCAAGGAGAAGCAGCTCTTGCTC  
GGTTGGGAGACGGTGCAAGAGAATCTGCCCCCTATAGGGGAATGGTGCGCACAGCCCTAGGGATC  
ATTGAAGAGGAAGGCTTTCTAAAGCTTTGGCAAGGAGTGACACCCGCCATTTACAGACACGTAGT  
GTATTCTGGAGGTCGAATGGTCACATATGAACATCTCCGAGAGGTTGTGTTTGGCAAAAGTGAAG  
ATGAGCATTATCCCCTTTGGAAATCAGTCATTGGAGGGATGATGGCTGGTGTATTGGCCAGTTT  
TTAGCCAATCCAACCTGACCTAGTGAAGGTTTCAAGATGCAAATGGAAGGAAAAGGAAACTGGAAGG  
AAAACCATTGCGATTTTCGTGGTGTACATCATGCATTTGCAAAAATCTTAGCTGAAGGAGGAATAC  
GAGGGCTTTGGGCAGGCTGGGTACCCAATATACAAAGAGCAGCACTGGTGAATATGGGAGATTTA  
ACCACTTATGATACAGTGAAACACTACTTGGTATTGAATACACCACTTGAGGACAATATCATGAC  
TCACGGTTTATCAAGTTTATGTTCTGGACTGGTAGCTTCTATTCTGGGAACACCAGCCGATGTCA  
TCAAAAGCAGAATAATGAATCAACCACGAGATAAACAAGGAAGGGGACTTTTGTATAAATCATCG  
ACTGACTGCTTGATTCAAGGCTGTTCAAGGTGAAGGATTCATGAGTCTATATAAAGGCTTTTTACC  
ATCTTGGCTGAGAATGACCCCTTGGTCAATGGTGTCTGGCTTACTTATGAAAAAATCAGAGAGA  
TGAGTGGAGTCAGTCCATTT**TAA**

126/168

## **FIGURE 126**

MSVPEEEERLLPLTQRWPRASKFLLSGCAATVAELATFPLDLTKTRLQMGEAALARLGDGARES  
APYRGMVRTALGIIEEGFLKLWQGVTPAIYRHVVYSGGRMVITYEHLREVVFVGKSEDEHYPLWKS  
VIGGMMAGVIGQFLANPTDLVKVQMOMEGKRKLEGKPLRFRGVHHAFAKILAEGGIRGLWAGWVP  
NIQRAALVNMGDLTTYDTVKHYLVNTPLEDNIMTHGLSSLCSGLVASILGTPADVIKSRIMNQP  
RDKQGRGLLYKSSTDCLIQAVQGEFMSLYKGFLPSWLRMTPWSMVFWLTYEKIREMSGVSPF

### **Transmembrane domains:**

amino acids 25-38, 130-147, 233-248

**FIGURE 127**

CGCGGATCGGACCCAAGCAGGTCGGCGGGCGGCGGCAGGAGAGCGGCCGGGCGTCAGCTCCTCGAC  
CCCCGTGTCGGGCTAGTCCAGCGAGGCGGACGGGCGGCGTGGGCCCATGGCCAGGCCCGGCATGG  
AGCGGTGGCGCGACCGGCTGGCGCTGGTGACGGGGGCCCTCGGGGGGCATCGGCGCGGCCGTGGCC  
CGGGCCCTGGTCCAGCAGGGACTGAAGGTGGTGGGCTGCGCCCGCACTGTGGGCAACATCGAGGA  
GCTGGCTGCTGAATGTAAGAGTGCAGGCTACCCCGGGACTTTGATCCCCTACAGATGTGACCTAT  
CAAATGAAGAGGACATCCTCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGCGGTGTAGACATC  
TGCATCAACAATGCTGGCTTGGCCCGGCCTGACACCCTGCTCTCAGGCAGCACCAAGTGGTTGGAA  
GGACATGTTCAATGTGAACGTGCTGGCCCTCAGCATCTGCACACGGGAAGCCTACCAGTCCATGA  
AGGAGCGGAATGTGGACGATGGGCACATCATTAACATCAATAGCATGTCTGGCCACCGAGTGTTA  
CCCCTGTCTGTGACCCACTTCTATAGTGCCACCAAGTATGCCGTCACTGCGCTGACAGAGGGACT  
GAGGCAAGAGCTTCGGGAGGCCCAGACCCACATCCGAGCCACGTGCATCTCTCCAGGTGTGGTGG  
AGACACAATTGCGCTTCAAACCTCCACGACAAGGACCCTGAGAAGGCAGCTGCCACCTATGAGCAA  
ATGAAGTGTCTCAAACCCGAGGATGTGGCCGAGGCTGTTATCTACGTCCTCAGCACCCCGCACA  
CATCCAGATTGGAGACATCCAGATGAGGCCCACGGAGCAGGTGACCTAGTGACTGTGGGAGCTCC  
TCCTTCCCTCCCCACCCTTCATGGCTTGCCTCCTGCCTCTGGATTTTAGGTGTTGATTTCTGGAT  
CACGGGATACCACTTCCTGTCCACACCCCGACCAGGGGCTAGAAAATTTGTTTGAGATTTTATA  
TCATCTTGTCAAATTGCTTCAGTTGTAAATGTGAAAAATGGGCTGGGGAAAGGAGGTGGTGTCCC  
TAATTGTTTTACTTGTTAACTTGTTCTTGCGCCCTGGGCACTTGGCCTTTGTCTGCTCTCAGTG  
TCTTCCCTTTGACATGGGAAAGGAGTTGTGGCCAAAATCCCCATCTTCTTGACCTCAACGTCTG  
TGGCTCAGGGCTGGGGTGGCAGAGGGAGGCCTTCACCTTATATCTGTGTTGTTATCCAGGGCTCC  
AGACTTCCTCCTCTGCCTGCCCCACTGCACCCTCTCCCCCTTATCTATCTCCTTCTCGGCTCCCC  
AGCCCAGTCTTGGCTTCTTGTCCCCTCCTGGGGTCATCCCTCCACTCTGACTCTGACTATGGCAG  
CAGAACACCAGGGCCTGGCCCAGTGGATTTTCATGGTGATCATTAAGAAAGAAAAATCGCAACCAA  
AAAAAAAAAA

## **FIGURE 128**

MARPGMERWRDRLALVTGASGGIGA AAVARALVQQGLKVVG CARTVGNIEELAAECKSAGYPGTLI  
PYRCDLSNEEDILSMFSAIRSQHSGVDICINNAGLARPD TLLSGSTSGWKDMFNVNVLALSICTR  
EAYQSMKERNVDDGHIININSMGHRVLP LSVTHFY SATKYAVTALTEGLRQELREAQTHIRATC  
ISPGVVETQFAFKLHDKDPEKAAATYEQMKCLKPEDVAEAVIYVLSTPAHIQIGDIQMRPTEQVT

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-17

#### **N-myristoylation sites.**

amino acids 18-24, 21-27, 22-28, 24-30, 40-46, 90-96, 109-115, 199-205

#### **Short-chain alcohol dehydrogenase.**

amino acids 30-42, 104-114

**FIGURE 129**

AACTTCTACATGGGCCTCCTGCTGCTGGTGCTCTTCCTCAGCCTCCTGCCGGTGGCCTACACCAT  
CATGTCCCTCCCACCCTCCTTTGACTGCGGGCCGTTCAAGGTGCAGAGTCTCAGTTGCCCGGGAGC  
ACCTCCCCCTCCCGAGGCAGTCTGCTCAGAGGGCCTCGGCCCAGAATTCCAGTTCTGGTTTCATGC  
CAGCCTGTAAAAGGCCATGGAACTTTGGGTGAATCACCGATGCCATTTAAGAGGGTTTTCTGCCA  
GGATGGAAATGTTAGGTCGTTCTGTGTCTGCGCTGTTCAATTCAGTAGCCACCAGCCACCTGTGG  
CCGTTGAGTGCTTGAAATTGAGGAACTGAGAAAATTAATTTCTCATGTATTTTTCTCATTTATTTA  
TTAATTTTTTAAGTATAGTTGTACATATTTGGGGGTACATGTGATATTTGGATACATGTATACAA  
TATATAATGATCAAATCAGGGTAACTGGGATATCCATCACATCAAACATTTATTTTTTATTCTTT  
TTAGACAGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAGTGGTGCCATCTCAGCTTACTGCAAC  
CTCTGCCTGCCAGGTTCAAGCGATTCTCATGCCTCCACCTCCCAAGTAGCTGGGACTACAGGCAT  
GCACCACAATGCCCAACTAATTTTTGTATTTTTAGTAGAGACGGGGTTTTGCCATGTTGCCCAGG  
CTGGCCTTGAACTCCTGGCCTCAAACAATCCACTTGCCTCGGCCTCCCAAAGTGTTATGATTACA  
GGCGTGAGCCACCGTGCCTGGCCTAAACATTTATCTTTCTTTGTGTTGGGAACTTTGAAATTAT  
ACAATGAATTATTGTAACTGTCATCTCCCTGCTGTGCTATGGAACACTGGGACTTCTTCCCTCT  
ATCTAACTGTATATTTGTACCAGTTAACCAACCGTACTTCATCCCCACTCCTCTCTATCCTTCCC  
AACCTCTGATCACCTCATTCTACTCTCTACCTCCATGAGATCCACTTTTTTAGCTCCCACATGTG  
AGTAAGAAAATGCAATATTTGTCTTTCTGTGCCTGGCTTATTTCACTTAACATAATGACTTCCTG  
TTCCATCCATGTTGCTGCAAATGACAGGATTTTCGTTCTTAATTTCAATTAAAATAACCACACATG  
GCAAAAA

130/168

## **FIGURE 130**

MGLLLLVLFLSLLPVAYTIMSLPPSFDCGPFRCRVSVAREHLPSRGSLLRGPRPRI PVLVSCQPV  
KGHGTLGESPMPEFKRVFCQDGNVRSFCVCAVHFSSHQPPVAVECLK

**Important features of the protein:**

**Signal peptide:**

amino acids 1-18

**N-myristoylation site.**

amino acids 86-92

**Zinc carboxypeptidases, zinc-binding region 2 signature.**

amino acids 68-79

**FIGURE 131**

TTCTGAAGTAACGGAAGCTACCTTGTATAAAGACCTCAACACTGCTGACCATGATCAGCGCAGCC  
 TGGAGCATCTTCCTCATCGGGACTAAAATTGGGCTGTTCCCTTCAAGTAGCACCTCTATCAGTTAT  
 GGCTAAATCCTGTCCATCTGTGTGTCGCTGCGATGCGGGTTTCATTTACTGTAATGATCGCTTTC  
 TGACATCCATTCCAACAGGAATACCAGAGGATGCTACAACCTCTCTACCTTCAGAACAAACAAATA  
 AATAATGCTGGGATTCCCTCAGATTTGAAAACTTGCTGAAAGTAGAAAGAATATACCTATACCA  
 CAACAGTTTAGATGAATTTCCCTACCAACCTCCCAAAGTATGTAAAAGAGTTACATTTGCAAGAAA  
 ATAACATAAGGACTATCACTTATGATTCACTTTCAAAAATTCCCTATCTGGAAGAATTACATTTA  
 GATGACAACCTCTGTCTCTGCAGTTAGCATAGAAGAGGGAGCATTCGAGACAGCAACTATCTCCG  
 ACTGCTTTTCCTGTCCCGTAATCACCTTAGCACAATTCCTTGGGGTTTGCCAGGACTATAGAAG  
 AACTACGCTTGGATGATAATCGCATATCCACTATTTTCATCACCATCTCTTCAAGGTCTCACTAGT  
 CTA AACCGCCTGGTTCTAGATGGAAACCTGTTGAACAATCATGGTTTAGGTGACAAAGTTTTCTT  
 CAACCTAGTTAATTTGACAGAGCTGTCCCTGGTGC GGAATTCCCTGACTGCTGCACCAGTAAACC  
 TTCCAGGCACAAACCTGAGGAAGCTTTATCTTCAAGATAACCACATCAATCGGGTGCCCCCAAAT  
 GCTTTTTTCTTATCTAAGGCAGCTCTATCGACTGGATATGTCCAATAATAACCTAAGTAATTTACC  
 TCAGGGTATCTTTGATGATTTGGACAATATAACACAACCTGATTCTTCGCAACAATCCCTGGTATT  
 GCGGGTGCAAGATGAAATGGGTACGTGACTGGTTACAATCACTACCTGTGAAGGTCAACGTGCGT  
 GGGCTCATGTGCCAAGCCCCAGAAAAGGTTTCGTGGGATGGCTATTAAGGATCTCAATGCAGAACT  
 GTTTGATTGTAAGGACAGTGGGATTGTAAGCACCATTCAAGATAACCACTGCAATACCCAACACAG  
 TGTATCCTGCCAAGGACAGTGGCCAGCTCCAGTGACCAAACAGCCAGATATTAAGAACCCCAAG  
 CTCACTAAGGATCAACAAACCACAGGGAGTCCCTCAAGAAAAACAATTACAATTACTGTGAAGTC  
 TGTACCTCTGATACCATTATATCTCTTGGAACTTGCTCTACCTATGACTGCTTTGAGACTCA  
 GCTGGCTTAACTGGGCCATAGCCCGGCATTTGGATCTATAACAGAAACAATTGTAACAGGGGAA  
 CGCAGTGAGTACTTGGTCACAGCCCTGGAGCCTGATTACCCCTATAAAGTATGCATGGTTCCCAT  
 GGAAACCAGCAACCTCTACCTATTTGATGAACTCCTGTTTGTATTGAGACTGAACTGCACCCC  
 TTCGAATGTACAACCCTACAACCACCCTCAATCGAGAGCAAGAGAAAGAACCTTACAAAAACCCC  
 AATTTACCTTTGGCTGCCATCATTTGGTGGGGCTGTGGCCCTGGTTACCATTGCCCTTCTTGCTTT  
 AGTGTGTTGGTATGTTTCATAGGAATGGATCGCTCTTCTCAAGGAAGTGTGCATATAGCAAAGGGA  
 GGAGAAGAAAGGATGACTATGCAGAAGCTGGCACTAAGAAGGACAACTCTATCCTGGAAATCAGG  
 GAACTTCTTTTCAGATGTTACCAATAAGCAATGAACCCATCTCGAAGGAGGAGTTTGTAATACA  
 CACCATATTTCCCTCCTAATGGAATGAATCTGTACAAAACAATCACAGTGAAAGCAGTAGTAACC  
 GAAGCTACAGAGACAGTGGTATTCCAGACTCAGATCACTCACACTCATGATGCTGAAGGACTCAC  
 AGCAGACTTGTGTTTTGGGTTTTTTAAACCTAAGGGAGGTGATGGT

## **FIGURE 132**

MISAAWSIFLIGTKIGLFLQVAPLSVMAKSCPSVCRC DAGFIYCND RFLTSIPTGIPEDATTLYL  
 QNNQINNAGIPSDLKNLLKVERIYLYHNSLDEFPTNL PKYVKELHLQENNI RTITYDSLSKIPYL  
 EELHLDDNSVSAVSIEEGAFRDSNYLRLLFLSRNHLSTIPWGLPRTIEELRLDDNRISTISSPSL  
 QGLTSLKRLVLDGNLLNNHGLGDKVFFNLVNLTELSLVRNSLTAAPVNLPGTNLRKLYLQDNHIN  
 RVPPNAFSYLRQLYRLDMSNNNLSNLPQGI FDDLDNITQLILRNNPWYCGCKMKWVRDWLQSLPV  
 KVNVRGLMCQAPEKVRGMAIKDLNAELFDCKDSGIVSTIQITTAIPNTVYPAQGQWPAPVTKQPD  
 IKNPKLTKDQQTGSPSRKTITITVKS VTSDTIHISWKLALPMTALRLSWLKLGHSPAFGSITET  
 IVTGERSEYLVTALEPDSPYKVCMPMETS NLYLFDETPVCIETETAPLRMYNPTTTLNREQEKE  
 PYKNPNLPLAAIIGGAVALVTIAL LALVCWYVHRNGSLFSRNCAYSKGRRRKDDYAEAGTKKDNS  
 ILEIRETSFQMLPISNEPISKEEFVIHTIFPPNGMNLYKNNHSESSSNRSYRDSGIPDSDHSHS

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-28

#### **Transmembrane domain:**

amino acids 531-552

#### **N-glycosylation sites.**

amino acids 226-229, 282-285, 296-299, 555-558, 626-629, 633-636

#### **Tyrosine kinase phosphorylation site.**

amino acids 515-522

#### **N-myristoylation sites.**

amino acids 12-17, 172-177, 208-213, 359-364, 534-539, 556-561,  
 640-645

#### **Amidation site.**

amino acids 567-570

#### **Leucine zipper pattern.**

amino acids 159-180

#### **Phospholipase A2 aspartic acid active site.**

amino acids 34-44



**FIGURE 133**

CCGTCATCCCCCTGCAGCCACCCTTCCCAGAGTCCTTTGCCCAGGCCACCCCAGGCTTCTTGGA  
GCCCTGCCGGGCCACTTGTCTT**CATG**TCTGCCAGGGGGAGGTGGGAAGGAGGTGGGAGGAGGGCG  
TGCAGAGGCAGTCTGGGCTTGCCAGAGCTCAGGGTGCTGAGCGTGTGACCAGCAGTGAGCAGAG  
GCCGGCCATGGCCAGCCTGGGGCTGCTGCTCCTGCTCTTACTGACAGCACTGCCACCGCTGTGGT  
CCTCCTCACTGCCTGGGCTGGACACTGCTGAAAGTAAAGCCACCATTGCAGACCTGATCCTGTCT  
GCGCTGGAGAGAGCCACCGTCTTCCTAGAACAGAGGCTGCCTGAAATCAACCTGGATGGCATGGT  
GGGGGTCCGAGTGCTGGAAGAGCAGCTAAAAAGTGTCCGGGAGAAGTGGGCCCAGGAGCCCCCTGC  
TGCAGCCGCTGAGCCTGCGCGTGGGGATGCTGGGGGAGAAGCTGGAGGCTGCCATCCAGAGATCC  
CTCCACTACCTCAAGCTGAGTGATCCCAAGTACCTAAGAGAGTTCCAGCTGACCCTCCAGCCCGG  
GTTTTGGAAGCTCCACATGCCTGGATCCACACTGATGCCTCCTTGGTGTACCCACGTTTCGGGC  
CCCAGGACTCATTCTCAGAGGAGAGAAGTGACGTGTGCCTGGTGCAGCTGCTGGGAACCGGGACG  
GACAGCAGCGAGCCCTGCGGCCTCTCAGACCTCTGCAGGAGCCTCATGACCAAGCCCGGCTGCTC  
AGGCTACTGCCTGTCCCACCAACTGCTCTTCTTCCTCTGGGCCAGAATGAGGGGATGCACACAGG  
GACCACTCCAACAGAGCCAGGACTATATCAACCTCTTCTGCGCCAACATGATGGACTTGAACCGC  
AGAGCTGAGGCCATCGGATACGCCTACCCTACCCGGGACATCTTCATGGAAAACATCATGTTCTG  
TGGAATGGGCGGCTTCTCCGACTTCTACAAGCTCCGGTGGCTGGAGGCCATTCTCAGCTGGCAGA  
AACAGCAGGAAGGATGCTTCGGGGAGCCTGATGCTGAAGATGAAGAATTATCTAAAGCTATTCAA  
TATCAGCAGCATTTTTTCGAGGAGAGTGAAGAGGCGAGAAAAACAATTTCCAGATTCTCGCTCTGT  
TGCTCAGGCTGGAGTACAGTGGCGCAATCTCGGCTCACTGCAACCTTTGCCTCCTGGGTTCAGC  
AATTCTCTTGCTCATCCTCCCGAGTAGCTGGGACTACAGGAGCGTGCCACCATACTGGCTAAT  
TTTTATATTTTTTTAGTAGAGACAGGGTTTCATCATGTTGCTCATGCTGGTCTCGAACTCCTGAT  
CTCAAGAGATCCGCCCACCTCAGGCTCCCAAAGTGTGGGATTAT**TAGG**TGTGAGCCACCGTGTCTG  
GCTGAAAAGCACTTTCAAAGAGACTGTGTTGAATAAAGGGCCAAGGTTCTTGCCACCCAGCACTC  
ATGGGGGCTCTCTCCCCTAGATGGCTGCTCCTCCCACAACACAGCCACAGCAGTGGCAGCCCTGG  
GTGGCTTCCTATACATCCTGGCAGAATAACCCCCAGCAAACAGAGAGCCACACCCATCCACACCG  
CCACCACCAAGCAGCCGCTGAGACGGACGGTTCCATGCCAGCTGCCTGGAGGAGGAACAGACCCC  
TTTAGTCCTCATCCCTTAGATCCTGGAGGGCACGGATCACATCCTGGGAAGAAGGCATCTGGAGG  
ATAAGCAAAGCCACCCGACACCCAATCTTGGAAGCCCTGAGTAGGCAGGGCCAGGGTAGGTGGG  
GGCCGGGAGGGACCCAGGTGTGAACGGATGAATAAAGTTCAACTGCAACTGAAAAAAAAAAAA

## **FIGURE 134**

MSARGRWEGGGRRACRGSGLLARAQGAERVTSSEQRPAMASLGLLLLLLLLLTALPPLWSSSLPGLD  
TAESKATIADLILSALERATVFLEQRLPEINLDGMVGVRVLEEQLKSVREKWAQEPLLQPLSLRV  
GMLGEKLEAAIQRSLHYLKLSDPKYLREFQLTLQPGFWKLPHAWIHTDASLVYPTFGPQDSFSEE  
RSDVCLVQLLGTGTDSSSEPCGLSDLCRSLMTKPGCSGYCLSHQLLFFLWARMRGCTQGPLQQSQD  
YINLFCANMMDLNRRAEAIGYAYPTRDIFMENIMFCGMGGFSDFYKLRWLEAILSQKQQEGCFG  
EPDAEDEELSKAIQYQQHFSTRVKRREKQFPDSRVAQAGVQWRNLGSLQPLPPGFKQFSCILIP  
SSWDYRSVPPYLANFYIFLVETGFHHVAHAGLELLISRDPPTSGSQSVGL

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-26

#### **Transmembrane domain:**

amino acids 39-56

#### **Tyrosine kinase phosphorylation sites.**

amino acids 149-156, 274-282

#### **N-myristoylation sites.**

amino acids 10-16, 20-26, 63-69, 208-214

#### **Amidation site.**

amino acids 10-14

#### **Glycoprotein hormones beta chain signature 1.**

amino acids 230-237

**FIGURE 135**

GGTCTGAGTGCAGAGCTGCTGTC**ATG**GCGGCCGCTCTGTGGGGCTTCTTTCCCGTCCTGCTGCTG  
CTGCTGCTATCGGGGGATGTCCAGAGCTCGGAGGTGCCCGGGGCTGCTGCTGAGGGATCGGGAGG  
GAGTGGGGTCGGCATAGGAGATCGCTTCAAGATTGAGGGGCGTGCAGTTGTTCCAGGGGTGAAGC  
CTCAGGACTGGATCTCGGCGGCCCGAGTGCTGGTAGACGGAGAAGAGCACGTCGGTTTCCTTAAG  
ACAGATGGGAGTTTTGTGGTTCATGATATACCTTCTGGATCTTATGTAGTGGAAGTTGTATCTCC  
AGCTTACAGATTTGATCCCGTTTCGAGTGGATATCACTTCGAAAGGAAAAATGAGAGCAAGATATG  
TGAATTACATCAAAACATCAGAGGTTGTCAGACTGCCCTATCCTCTCCAAATGAAATCTTCAGGT  
CCACCTTCTTACTTTATTAAAAGGGAATCGTGGGGCTGGACAGACTTTCTAATGAACCAATGGT  
TATGATGATGGTTCTTCCTTTATTGATATTTGTGCTTCTGCCTAAAGTGGTCAACACAAGTGATC  
CTGACATGAGACGGGAAATGGAGCAGTCAATGAATATGCTGAATTCCAACCATGAGTTGCCTGAT  
GTTTCTGAGTTCATGACAAGACTCTTCTCTTCAAATCATCTGGCAAATCTAGCAGCGGCAGCAG  
TAAACAGGCAAAAGTGGGGCTGGCAAAAGGAGG**TAG**TCAGGCCGTCCAGAGCTGGCATTTCAC  
AAACACGGCAACACTGGGTGGCATCCAAGTCTTGAAAACCGTGTGAAGCAACTACTATAAACTT  
GAGTCATCCCGACGTTGATCTCTTACAACGTGTGTATGTT  
AACTTTTTAGCACATGTTTTGTACTTGGTACACGAGAAAACCCAGCTTTCATCTTTTGTCTGTAT  
GAGGTCAATATTGATGTCACTGAATTAATTACAGTGTCTATAGAAAATGCCATTAATAAATTAT  
ATGAACTACTATACATTATGTATATTAATTAACATCTTAATCCAGAAATCAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAA